ATMOSPHERIC EXCITATION OF POLAR MOTION

by

John William Kuehne, B.M., B.S., M.A.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

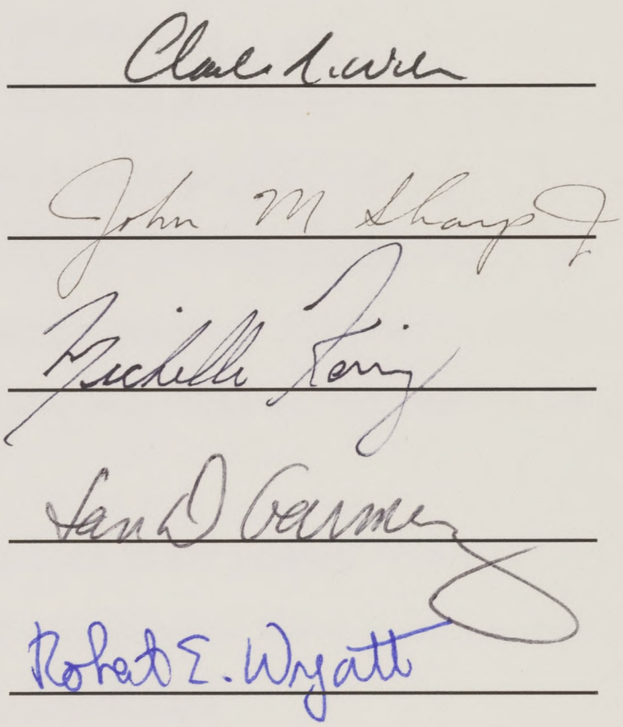
The University of Texas at Austin

August 1996

Atmospheric Excitation of Polar Motion

Approved by

Dissertation Committee:



Dedicated to my Father, Robert A. Kuehne

Preface

Whatever Fortune brings, don't be afraid of doing things. (especially, of course, for Kings) -- A.A. Milne, *Now We Are Six*

This dissertation comes from many people in different ways. First is my supervisor, Clark Wilson, who made this research possible. Besides generous funding leading to a (probable) department record for student tenure and travel, Clark has allowed me to work unrestricted, and yet his knowledge, insight, and intuition into the polar motion problem has always guided this research.

Stuart Johnson also figures prominently. Clark assigned him as an undergraduate assistant to this project, but unlike previous assistants he plowed through a pile of polar motion papers, and quickly left me feeling like the undergraduate. He lunged at the problem with verve and enthusiasm, making a number of important contributions to the 1993 paper. The time that I spent working with Stuart was the most rewarding and fun part of this study.

My wife, Deborah Swart, while being a pillar in our personal relationship and growth, has also been an important scientific co-supervisor. Much encouragement to try something new or look at some data came from Debbie.

This journey started during the Spring of 1984, when my father and I drove from Lexington, Kentucky to Austin to visit Hildegard and Rizer Everett (Hildegard and my father are cousins). Their family has long been rooted in Texas, and he not-so-secretly wanted me to attend graduate school somewhere in Texas. My father did not live past 1984 to see me here at UT, but I think from our last conversations that he knew I was headed for Texas.

Atmospheric Excitation of Polar Motion

Publication No. _____

John William Kuehne, Ph.D.
The University of Texas at Austin, 1996

Supervisor: Clark R. Wilson

Analysis of nonseasonal polar motion excitation and atmospheric mass equatorial angular momentum (EAM) over land for the period 1980-1993 reveals a clear pattern of high power and correlation during the northern hemisphere (NH) winter followed by low power and correlation during the NH summer. During the southern hemisphere winter there is significant correlation between the atmospheric EAM over midlatitude southern oceans and polar motion excitation indicating the existence of a dynamic atmosphere-ocean excitation. The atmospheric excitation power is too small to explain the large correlation during the NH winter. The momentum of wind probably accounts for the deficit in power. The implication of these results is that there are two main excitation sources each dominant at different seasons. Atmospheric mass redistribution over land forces polar motion during the NH winter, and a dynamic atmosphere-ocean response is important during the SH winter.

The Chandler wobble frequency F and quality factor Q are estimated by least squares using polar motion data in combination with proxy excitations derived from atmospheric data. Monte Carlo tests show that this approach can provide useful improvements over traditional maximum likelihood methods, which use only observed polar motion. With less than a decade of good simultaneous polar motion and atmospheric data there is no new information about Q, but 90% confidence intervals for F are near ± 0.004 cycles per year (cpy), making them comparable to maximum likelihood results based on nearly a century of historical polar motion data. Using a proxy excitation constructed from a linear combination of atmospheric EAM over land and SH ocean, the estimate of F is near 0.831 cpy (a period of 439.5 days), which is significantly different from the maximum likelihood value near 0.843 cpy (433 days). This difference can be understood in light of the assumptions underlying the two methods.

Table of Contents

Introduction

One of the fundamental properties of the Earth which favors life is its rotation. Some other properties include atmosphere, oceans, tectonic activity, and distance from the Sun. The Earth must have attained its rapid rate of rotation early in its history. It is remarkable that all the planets except Venus and Mercury are rotating rapidly.

The orientation of the Earth in a celestial reference frame is now routinely monitored with high accuracy using space geodetic techniques. These data are used to infer the orientation of the rotation axis in the celestial frame (precession and nutation), its spin rate (length-of-day), and its location in the terrestrial reference frame (polar motion). Gravitational torques of the moon, sun, and planets acting on the Earth's equatorial bulge slowly change the orientation of the angular momentum vector and the attending rotation axis. The resulting precessions and nutations, while of considerable importance, are not the topic of this dissertation.

The motions of the atmosphere, ocean, and liquid core change the rotation of the Earth. The conservation of angular momentum is the guiding principle for understanding how the rotation changes when one part of the Earth moves relative to another. For example, considering only the solid Earth and its atmosphere, the conservation principle allows one to determine the change in total atmospheric angular momentum by observing the change in the Earth's rotation (and vice versa) without knowledge of how the momentum is transferred.

Polar motion was first observed nearly a century ago and meteorological phenomena were immediately suggested as the excitation source [*Munk and MacDonald*, 1960]. Aside from the annual motion (see appendix D), the

meteorological hypothesis was unproved until *Barnes, Hide, et al.* [1983] showed that part of polar motion could be explained by fluctuations in atmospheric mass angular momentum. In conjunction with highly accurate space-geodetic polar motion data, their discovery was made possible by the advent of global circulation models, which produce estimates of pressure and winds at regular time intervals. Since their paper, a number of reports have appeared which suggest that between 9% and 25% of the polar motion variance is linearly related to atmospheric mass angular momentum, most notably *Eubanks et al*, [1988]. To date there is no evidence that seismic events, or any solid-Earth mechanism, significantly excites polar motion.

The explanation for the remaining polar motion variance is some combination of the following:

Winds. The total atmospheric angular momentum is composed of two parts. The mass (or pressure) term is that part of the momentum which changes the inertia of the Earth, including its atmosphere. The mass term is estimated from surface pressure data, which is perhaps the best determined atmospheric parameter. The motion (or wind) term is the angular momentum of motion relative to the solid Earth. The wind term is difficult to estimate because it requires knowledge of the atmospheric velocity and density throughout the total volume of the atmosphere.

Oceans. The atmospheric mass term is estimated by assuming that the oceans compensate local atmospheric pressure fluctuations, so that pressure fluctuations on the sea floor depend only on fluctuations of the mean atmospheric pressure over the entire ocean (the inverted barometer hypothesis). This convenient assumption probably is not true for periods less than about 5 days. The winds drive currents and change the pressure at the sea floor, but there are few current and pressure data available.

Water storage variations. Variations is terrestrial water storage perturb the inertia of the Earth, but the details are unknown. Water storage variations cannot explain rapid polar motions, and may be important only at periods longer than a year (*Kuehne and Wilson*, 1991).

Investigations in these categories have been hampered by lack of atmospheric, oceanic, and hydrologic data. Atmospheric modeling has been the most successful because the atmosphere is accessible and relevant to humans. Ocean models are driven by observations of surface winds, currents, and temperatures, which may be insufficient to model currents and pressure variations at the sea floor. Surprisingly, hydrologic variations are most difficult to model because there is no way to observe directly the amount of terrestrial water storage. The fundamental parameter - precipitation - is difficult to measure partly because it is not spatially correlated with itself. Hence good estimates of precipitation require a dense network of rain gauges. Evapotranspiration and runoff are also difficult to measure, and the required temporal integration of (precipitation-runoff-evapotranspiration) compounds the errors.

The oceans are the most promising target to investigate for two reasons. First, they cover most of the planet, so that even slight deviations from the inverted barometer hypothesis might produce significant polar motion. Second, currents are partly driven by the wind and pressure variations at the sea surface, and global circulation models provide estimates of these data.

Given that the oceans are partly driven by atmospheric processes which are observed and modeled, it is not entirely necessary to understand the physical processes occurring in the ocean to prove the hypothesis that the oceans cause polar motion. One can construct an empirical transfer function that transforms atmospheric data into polar motion for some time interval, and then test the success of the transfer

function for a different time interval. This is system identification. The transfer function which I proposed to use is a kind of generalized non-linear function often called a neural network.

Work on the neural network system identification showed that atmospheric angular momentum is repeatedly correlated with polar motion. However, this can be demonstrated directly by computing the correlation over short (150-day) overlapping windows of the time series. Such analysis of polar motion excitation and atmospheric mass equatorial angular momentum reveals a clear pattern of high power and correlation during the northern hemisphere (NH) winter followed by low power and correlation during the NH summer.

The windowed correlation analysis also supports the hypothesis that the oceans are important in polar motion excitation. During the southern hemisphere winter there is significant correlation between the atmospheric mass angular momentum over midlatitude southern oceans and polar motion excitation indicating the existence of a dynamic atmosphere-ocean excitation, or some other effect of the winds.

These fundamental and salient properties of the atmospheric and polar motion time series were not reported in the literature when this research began. With these discoveries, work on system identification ceased in favor of direct investigation into the excitation of polar motion. Appendix C contains a summary of the neural network system identification method used, and some examples.

This dissertation is organized in two major sections and three appendices. The first section is a published paper in the Journal of Geophysical Research, and appears as *Atmospheric excitation of non-seasonal polar motion* by John Kuehne, Stuart Johnson, and Clark R. Wilson, v.98, pp.19,973-19,978, 1993. This contains the fundamental results and interpretations of the research. The second section, which is

published in the Journal of Geophysical Research as *Estimation of the Chandler wobble frequency and Q* by John Kuehne, Clark R. Wilson, and Stuart Johnson, v.101, pp.13573-13,579, 1996, is an application of the results. The American Geophysical Union has granted permission for the use of this material for inclusion in this dissertation, including microfilm editions thereof.

Appendix A, which appears in the first published paper, describes the calculation of the inverted barometer ocean response. Appendix B describes some new atmospheric data available since the publication of the paper. Appendix C is a complete derivation of the neural network filter, with some synthetic examples. Appendix D presents the results for the annual, i.e. seasonal, component of polar motion. Theoretical treatment of the physics of Earth rotation can be found in *Munk and MacDonald* [1960] and *Eubanks* [1993].

Atmospheric Excitation of Nonseasonal Polar Motion

Abstract: Analysis of nonseasonal polar motion excitation and atmospheric mass equatorial angular momentum (EAM) over land for the period 1980-1989 reveals a clear pattern of high power and correlation during the northern hemisphere (NH) winter followed by low power and correlation during the NH summer. A special case of this pattern occurs for longer than 14 months (from January 1987 to March 1988) when the correlation throughout the NH summer remains statistically significant. During this epoch an average of 72% of the nonseasonal polar motion excitation power at frequencies between -30 and +12 cycles/yr is linearly related to atmospheric EAM over land. During the southern hemisphere winter there is significant correlation between the atmospheric EAM over midlatitude southern oceans and polar motion excitation indicating the existence of a dynamic atmosphere-ocean excitation. The atmospheric excitation power is too small to explain the large correlation during the NH winter. The effects of winds probably account for the deficit in power. The implication of these results is that there are two main excitation sources each dominant at different seasons. Atmospheric mass redistribution over land forces polar motion during the NH winter, and a dynamic atmosphere-ocean response is important during the SH winter.

1. Introduction

A century after the discovery of the annual and 14-month components of polar motion by Chandler in 1892, the geophysical processes which excite polar motion at all frequencies remain enigmatic. Movement of air and water must be the excitation source of annual polar motion, although the details remain uncertain [*Chao and Au*, 1991]. At nonseasonal frequencies the variance of observed polar motion excitation exceeds the variance of atmospheric excitation, but significant correlation is observed [*Eubanks et al.*, 1988].

We analyze polar motion data and accompanying estimates of polar motion excitation derived from meteorological data for the period 1980-1989. This is the decade in which two remarkable observational accomplishments became routine: space-geodetic determination of polar motion at periods of 5 days and less, with accuracy eventually exceeding 1 milliarc second (mas); and global determinations of atmospheric variations, providing values of various atmospheric parameters on a global grid every few hours. Our analysis of these data differs from that of previous investigations [e.g., *Eubanks et al.*, 1988; *Chao and Au*, 1991; *Preisig,* 1992] mainly in the use of sliding-window estimates of power, correlation, and phase.

We illustrate a number of important features about atmospheric excitation of polar motion, including the general nonstationary nature of the time series, a seasonal variation in power and correlation, evidence for a nonequilibrium ocean excitation, and a 14-month period during which 72% of the polar motion variance at frequencies from -30 to +12 cycles per year (cpy) is related to atmospheric mass redistribution over land. These results demonstrate that during the study period, polar motion is principally excited by the motion of air and water.

2. Excitation Functions

Inferred Excitation from Polar Motion Data

Conservation of angular momentum relates two quantities evaluated in the rotating geographic (or Earth-fixed) coordinate system: deviations in the angular momentum of the Earth (including atmosphere) and pole position, described by the complex functions $\chi(t)$ and $m(t)$, respectively. The geographic coordinate system is defined by mutually orthogonal basis vectors e_1, e_2, and e_3. e_1 and e_2 are in the equatorial plane with e_1 intersecting the Greenwich meridian and e_2 intersecting the meridian 90° east. e_3 intersects the geographic north pole. Quantities with subscripts 1, 2, and 3 are components along these axes. The linearized Liouville equation is [*Gross*, 1992]

$$\chi(t) = \frac{i}{\sigma_c}\frac{dm(t)}{dt} + m(t) \qquad 1$$

where σ_c is the complex Chandler frequency (defined below), and $m(t) = m_1(t) + im_2(t)$ is the complex representation of the direction cosines reported in polar motion time series. Following *Barnes et al.* [1983],

$$\begin{aligned}\chi(t) &= \chi_1(t) + i\chi_2(t) \\ &= (1.00\Omega\Delta I(t) + 1.43h(t)) / \Omega(C - A)\end{aligned} \qquad 2$$

where $\Delta I(t) = \Delta I_{13}(t) + i\Delta I_{23}(t)$ is the complex deviation in the equatorial products of the Earth's inertia, $C = 7.04 \times 10^{37}\,\mathrm{kg\,m^2}$ is the principal moment of inertia, A is the

moment of inertia about an equatorial axis such that $(C-A)=0.00333C$, $\Omega=0.7292115\times10^{-4}$ radian per sidereal second is the mean angular velocity, the factors 1.00 and 1.43 account for surface loading and rotational deformation, and $h(t)=h_1(t)+ih_2(t)$ is the complex angular momentum due to motion in the geographic coordinate system. The $\Delta I(t)$ and $h(t)$ terms are evaluated from meteorological pressure and wind data, respectively.

We use a subset of a polar motion time series called Space90, which is described in the annual report of the International Earth Rotation Service for 1990. The data are a combined solution of satellite laser ranging and very long baseline interferometry observations interpolated to a daily time series for the period 1980-1989. The complex polar motion time series m_t are deconvolved to χ_t^m using the filter [*Wilson,* 1985]

$$\chi_t^m = \frac{ie^{-i\pi F_c T}}{\sigma_c T}\left(m_{t+T/2} - e^{i\sigma_c T} m_{t-T/2}\right), \qquad 3$$

which is a discrete form of (1). The complex Chandler frequency $\sigma_c = 2\pi F_c(1+i/2Q_c)$ is evaluated with F_c=0.843 cpy and Q_c=175 [*Wilson and Vicente,* 1980]. The filter introduces a shift in the sampling time of 1/2 the sample interval T, or 1/2 day. As the atmospheric data are available at 1/2-day intervals, this poses no problem in alignment. The filter (3) has an amplitude error with respect to (1), which causes high frequencies in the resulting χ_t^m to be reduced in amplitude by a factor which approaches 0.6 at the Nyquist. This results in an insignificant broadband attenuation of 0.25% in power because the polar motion time series has no

significant energy above approximately 1/10 cycles per day (cpd), well below the filter application Nyquist frequency of 1/2 cpd.

Estimated Atmospheric Excitation Functions

The atmospheric pressure data consist of twice daily 1000 and 850 mbar isobaric heights on a 2.5° grid from the European Centre for Medium-Range Weather Forecasts (ECMWF). A surface pressure p_s is obtained at each 2.5 degree grid point by exponential interpolation of the form

$$p_s = 1000 \exp\left[\ln(850/1000)(h_{1000} - h)/(h_{1000} - h_{850})\right] \qquad 4$$

where h is the grid elevation and h_{1000}, h_{850} are the isobaric heights. This method of obtaining surface pressures is chosen as it yields almost identically the global χ values provided separately by the ECMWF. The required surface topography is supplied by the National Center for Atmospheric Research (NCAR), which also distributes the ECMWF data.

Computation of the χ from the pressure data follows *Barnes et al.* [1983], who calculate $\Delta I(t)$ in spherical coordinates. The resulting integrals over latitude, longitude, and radius are converted (using the hydrostatic relationship between the radial and pressure coordinates) to integrals over latitude, longitude and pressure. Approximating the atmosphere by a thin shell at a constant radius $R = 6.371 \times 10^6$ m, the equatorial angular momentum due to atmospheric mass is

$$\chi_t^P = -(20.63 \times 10^7)\frac{1.00R^4}{(C-A)g}\iint d\lambda\, d\phi\, p_s(\lambda,\phi,t)e^{i\lambda}\sin\phi\cos^2\phi \qquad 5$$

where ϕ is latitude, λ is longitude (such that λ increases eastward and is zero along the Greenwich meridian), $g = 9.8\,\mathrm{m/s^2}$ is the acceleration of gravity at the surface, 1.00 accounts for the effect of atmospheric loading, and the factor of 20.63×10^7 converts radians to mas. Equation (5) is typically evaluated as sums over the entire globe. However, in this study we obtain separate χ_t^P for various geographic regions, especially for the land χ_t^{PL}, ocean χ_t^{PO} and subsections of each.

The isostatic behavior of the combined ocean and atmosphere, sometimes called the inverted barometer (IB) hypothesis, is not added to χ_t^{PL} because we are interested in the separate contributions of the atmosphere over land and ocean. The IB term, computed from atmospheric data over land (see the appendix), is small compared to χ_t^{PL}, and its addition to χ_t^{PL} does not significantly change our results. Significant correlation between χ_t^m and χ_t^{PO} during the southern hemisphere winter is evidence that a non-IB response is important, as explained in section 4.

Because we do not have the gridded wind data to evaluate $h(t)$, we use the ECMWF global χ_t^W, computed from the zonal and meridional winds at pressure levels up to 10 mbar, again using the hydrostatic relationship to integrate over pressure. Daily evaluations of χ_t^W from 1986 to the end of 1989 are used in this study.

3. Time Series Analysis

The linear least squares estimates of the mean, trend, annual sinusoid, and semiannual sinusoid are subtracted from each time series to produce nonseasonal time series. The time series are resampled to a uniform 5-day increment after applying a zero-phase low-pass time domain filter with a cutoff at 0.8 of the Nyquist. The 5-day

increment is chosen because there are no reliable variations present in the polar motion data at periods less than 3-5 days.

Our method of time domain analysis is based upon sliding-window power, correlation, and phase estimates. This technique permits the analysis of time series whose statistical properties change through time. A window is a short segment of the time series which is prepared with zero mean and trend. Estimates of the variance σ_τ^2, the squared magnitude of the complex correlation coefficient ρ_τ^2, and phase θ_τ are then computed for two time series χ_t, ξ_t as

$$\sigma_\tau^2(\chi) = \frac{\sum_{t=L(\tau)}^{U(\tau)} \chi_t \chi_t^*}{N-1} \qquad \rho_\tau^2(\chi,\xi) = \frac{\left| \sum_{t=L(\tau)}^{U(\tau)} \chi_t \xi_t^* \right|^2}{\left[\sum_{t=L(\tau)}^{U(\tau)} \chi_t \chi_t^* \right] \left[\sum_{t=L(\tau)}^{U(\tau)} \xi_t \xi_t^* \right]} \qquad \theta_\tau(\chi,\xi) = \arg\left(\sum_{t=L(\tau)}^{U(\tau)} \chi_t \xi_t^* \right) \tag{6}$$

where the asterisk denotes complex conjugation, N is the number of 5-day data in the window, $L(\tau) = \tau - (N/2)$, $U(\tau) = \tau + (N/2) - 1$, and τ is the middle time of the window. $\rho_\tau^2 \in [0,1]$ is the estimated fraction of linearly related variance between two time series. If $\chi_t = \xi_t + \eta_t$, where η_t is a noise time series uncorrelated with ξ_t, then $\rho_\tau^2(\chi,\xi)$ is approximated by $\sigma_\tau^2(\xi) / \sigma_\tau^2(\chi)$ when $\rho_\tau^2(\chi,\xi)$, $\sigma_\tau^2(\xi)$, and $\sigma_\tau^2(\chi)$ are good estimates. The angle $\theta_\tau \in [-\pi, \pi]$ is a weighted average over the window of the longitudinal difference between two time series so that positive phase indicates that χ_t is east of ξ_t.

Confidence levels for ρ_τ^2 are estimated from $\rho_\tau^2(\zeta, \chi^m)$, where χ_t^m is the observed polar motion excitation time series and ζ_t is a synthetic Gaussian noise time series, for 100 independent instances of ζ_t. We find the level above which 1% of the

$\rho_\tau^2(\zeta,\chi^m)$ fall, or the 99% confidence level. This method is an improvement over the standard approach which assumes that both time series are Gaussian, because the polar motion excitation time series is not statistically stationary. For N=30 (a 150-day window) the probability that any one ρ_τ^2 will exceed 0.18 in the absence of true correlation is less than 1%.

Frequency domain estimates of correlation and phase are based on the multiple taper method [*Thompson,* 1982], which provides superior control of spectral bias and leakage. High-resolution spectral estimates X_{fk} are computed by weighting the data χ_t with orthogonal tapers, each with associated eigenvalue w_k, and computing the discrete Fourier transform of each resulting time series. The spectral estimates are then

$$\rho_f^2(\chi,\xi)=\frac{\left|\sum_{k=1}^{q} w_k \mathrm{X}_{fk}\Xi_{fk}^*\right|^2}{\left(\sum_{k=1}^{q} w_k \mathrm{X}_{fk}\mathrm{X}_{fk}^*\right)\left(\sum_{k=1}^{q} w_k \Xi_{fk}\Xi_{fk}^*\right)} \qquad \rho_f^2(\chi,\xi)=\frac{\left|\sum_{k=1}^{q} w_k \mathrm{X}_{fk}\Xi_{fk}^*\right|^2}{\left(\sum_{k=1}^{q} w_k \mathrm{X}_{fk}\mathrm{X}_{fk}^*\right)\left(\sum_{k=1}^{q} w_k \Xi_{fk}\Xi_{fk}^*\right)}$$

$$\theta_f(\chi,\xi)=\arg\left(\sum_{k=1}^{q} w_k \mathrm{X}_{fk}\Xi_{fk}^*\right) \qquad 7$$

The q number of tapers determines the trade-off between spectral resolution and statistical confidence. For the spectral estimates with eight tapers presented in section 4 the resolution at each frequency f in cycles per year (cpy) is [f-3.36, f+3.36]cpy. The 99% confidence levels for ρ_f^2 are computed in the same manner as for the time domain estimates. In the absence of true correlation the probability of ρ_f^2 exceeding 0.5 at any one frequency is less than 1%.

The interpretation of coherence is similar to the time domain correlation. The phase indicates temporal lag or lead between the time series as a function of frequency so that positive prograde phase means that χ_t lags ξ_t.

Spectral analysis of the complex time series produces a separation of prograde and retrograde ($e^{\pm i\omega t}$) components. This is especially appropriate for study of the Chandler wobble, which is a prograde normal mode. At other frequencies the separation may describe zonal prograde and retrograde motion that we observe in the atmospheric data.

4. Results

We evaluate two different versions of χ_t^{PL}. The first is an integration over all land. The second version $\chi_t^{PL'}$ omits data from the Antarctic. Because this region is located at latitudes where the excitation is insensitive to mass redistribution (equation (5)) and the total area is small, the omission should have little effect. However, $\chi_t^{PL'}$ produces consistently higher correlation with polar motion excitation while retaining the variance of χ_t^{PL}, suggesting that the omitted data are noisy. *Trenberth and Olson* [1988] compared estimates of atmospheric quantities from the ECMWF and the National Meteorological Center (NMC) and reported discrepancies around Antarctica. This may indicate that the gridded-values over Antarctica are unreliable, presumably because there are insufficient observations from this region.

Time Domain Estimates

Examination of χ_t^m and χ_t^{PL} using the statistics σ_τ^2 and ρ_τ^2 reveals a yearly recurrence of high power and correlation during the northern hemisphere winter

followed by low power and correlation during the summer (Figures 1 and 2). This result is insensitive to the window length N provided that it is shorter than 1 year. We use N=30 (150 days). During NH winters, $\rho_\tau^2(\chi^m,\chi^{PL})$ is always statistically significant, greatly exceeding the 99% confidence level of 0.18, while during the summer it is insignificant except during 1987.

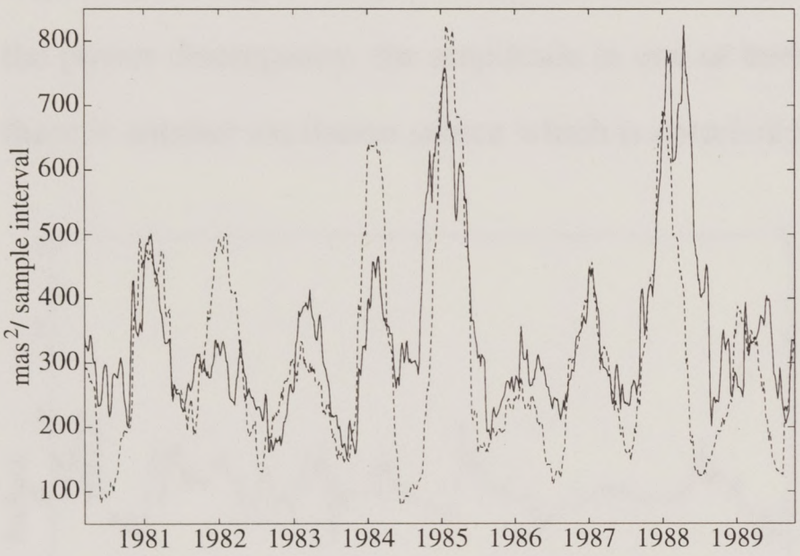


Fig. 1. The variance of the polar motion excitation $\sigma_\tau^2(\chi^m)/3$ (solid) and atmospheric mass angular momentum $\sigma_\tau^2(\chi^{PL})$ (dashed) estimated with a window of N=30 (150 days) centered on the dates shown. The scale factor of 1/3 is for visual comparison only.

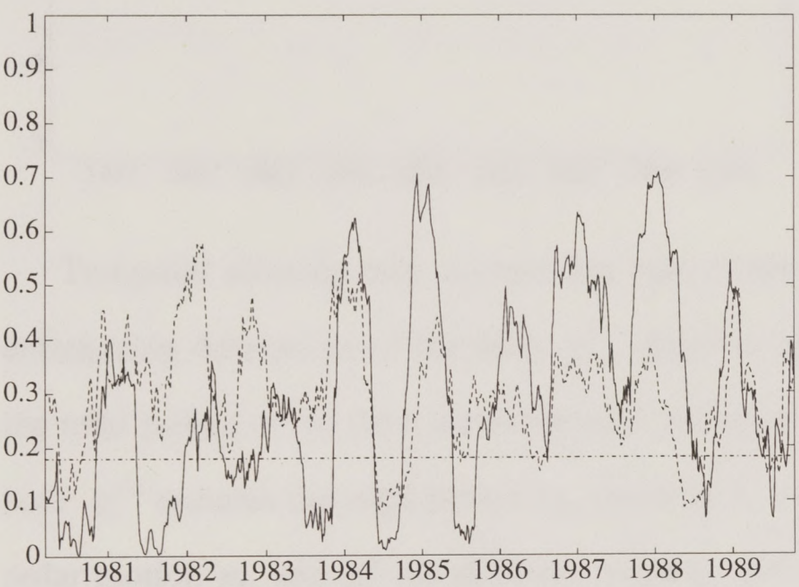


Fig. 2. Complex correlation squared $\rho_\tau^2(\chi^m,\chi^{PL})$ (solid) and the variance ratio $\sigma_\tau^2(\chi^{PL})/\sigma_\tau^2(\chi^m)$ (dashed) estimated with a window of N=30. The 99% confidence level refers to $\rho_\tau^2(\chi^m,\chi^{PL})$.

For several NH winters $\rho_\tau^2(\chi^m,\chi^{PL}) \cong 0.7$ (or a correlation coefficient of 0.84), implying that about 70% of $\sigma_\tau^2(\chi^{PL})$ is linearly related to $\sigma_\tau^2(\chi^m)$ at these times. The

phase is about zero when the correlation is significant except for isolated patches early in the time series (Figure 3). However, $\sigma_\tau^2(\chi^m)$ is larger than $\sigma_\tau^2(\chi^{PL})$ by a factor of about 2-3. In the absence of other excitation sources correlated with χ_t^{PL} we should expect that $\rho_\tau^2(\chi^m,\chi^{PL})$ is approximated by $\sigma_\tau^2(\chi^{PL})/\sigma_\tau^2(\chi^m)$. Although $\sigma_\tau^2(\chi^{PL})/\sigma_\tau^2(\chi^m)$ follows the yearly pattern of $\rho_\tau^2(\chi^m,\chi^{PL})$, it is usually smaller, especially during the larger correlation events. There are two obvious explanations for the power discrepancy: the amplitude in one or both of the time series is incorrect or there is another excitation source which is correlated with χ_t^{PL}.

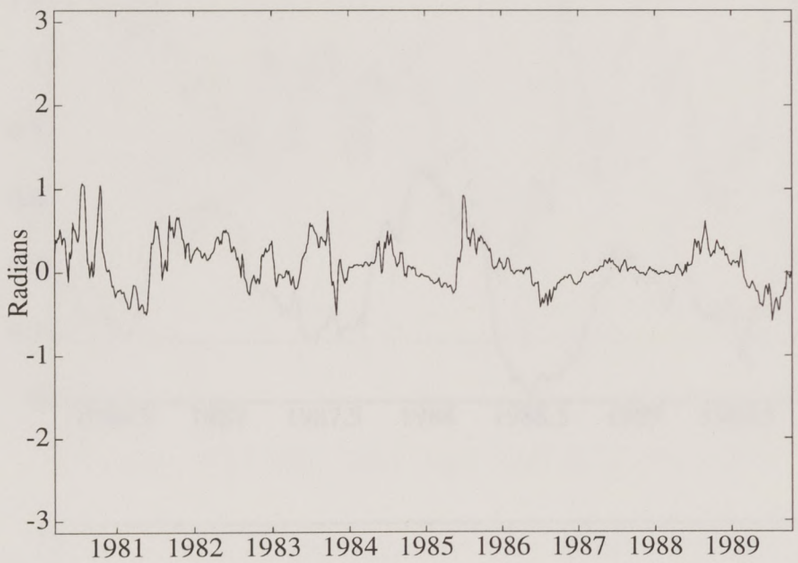


Fig. 3. The geographic phase of $\chi_t^{PL'}$ relative to χ_t^m. The phase remains small even when the correlation drops below the 99% confidence level.

Temporal aliasing may account for part of the power discrepancy. After 5-day antialiasing decimation of the daily χ_t^m, which is interpolated from 3- to -5 day data, the total power of the time series remains unchanged, whereas decimation of the 12-hour χ_t^{PL} reduces the total power by about 30%. However, aliasing of the observed polar motion m_t and χ_t^m are different because of the resonant amplification near the Chandler frequency. Using χ_t^{PL} as a model for χ_t^m, we estimate that less than 5% of

the power of χ_t^m is due to aliasing. Analysis of high-frequency polar motion data may provide a better estimate of the aliased power.

The wind contribution may also explain the power discrepancy. For some wintertime $\rho_\tau^2(\chi^m, \chi^{PL})$, $\sigma_\tau^2(\chi^{PL} + \chi^W)$ has ample power to explain polar motion while retaining most of the original correlation (Figure 4). The importance of the wind contribution is also reported by *Gross and Lindqwister* [1992].

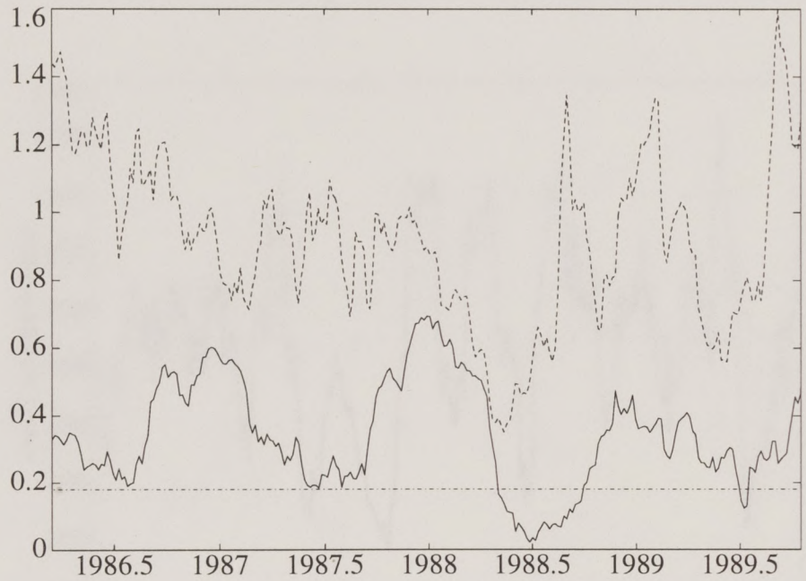


Fig. 4. Complex correlation squared $\sigma_\tau^2(\chi^{PL} + \chi^W)$ (solid) and the variance ratio $\sigma_\tau^2(\chi^{PL} + \chi^W) / \sigma_\tau^2(\chi^m)$ (dashed).

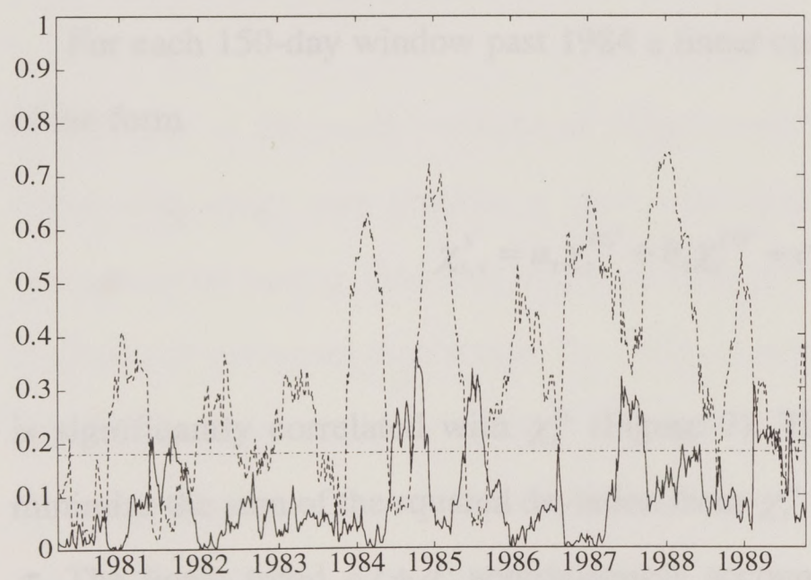


Fig. 5. Complex correlation squared $\rho_\tau^2(\chi^m, \chi^{PO'})$ (solid) and $\rho_\tau^2(\chi^m, \chi^{PL'})$ (dashed) for reference. Note that the ocean correlation peaks usually occur during the SH winter, 6 months out of phase with the land correlation peaks.

Usually during the SH winter, $\rho_\tau^2(\chi^m, \chi^{PO})$ is statistically significant. Virtually all of the correlation can be traced to the oceanic region roughly between 30° and 60° southern latitude (Figure 5). Between these latitudes, land covers less than 8% of the total area. Denoting this subregion with a prime, $\sigma_\tau^2(\chi^{PO'})$ also has a strong annual pattern of high power during the SH winter (Figure 6). When the correlation is significant, the phase averages so that $\chi_t^{PO'}$ is about 20° east of χ_t^m. We produce similar results with quantities computed from ocean surface wind data.

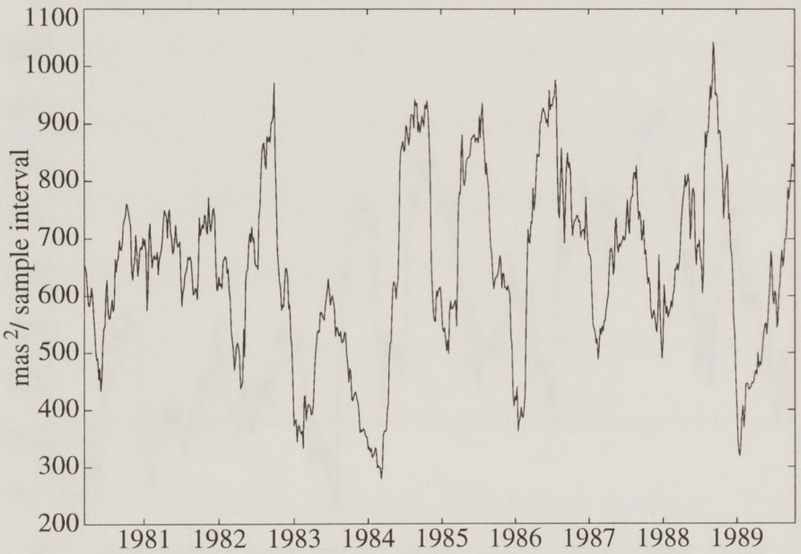


Fig. 6. The variance $\sigma_\tau^2(\chi^{PO'})$, computed from atmospheric mass variations over the SH oceans from 30° to 60° latitude. The variance is high during the SH winter.

For each 150-day window past 1984 a linear combination of $\chi_t^{PL'}$, $\chi_t^{PO'}$, t, and 1 of the form

$$\chi_{t,\tau}^S = a_\tau \chi_t^{PL'} + b_\tau \chi_t^{PO'} + c_\tau t + d_\tau \qquad 8$$

is significantly correlated with χ_t^m (Figure 7). The coefficients are computed to minimize the sum of the squared deviation from χ_t^m for each window centered at time τ. The linear trend $c_\tau t + d_\tau$ approximately accounts for variations longer than the 150-day period of analysis. Thus $\chi_{t,\tau}^S$ is the linear combination of $\chi_t^{PL'}$ and $\chi_t^{PO'}$

which best explains χ_t^m and maximizes the correlation during the 150-day period. Although the 99% confidence level increases from 0.18 to 0.23, $\rho_\tau^2(\chi^m,\chi_{,\tau}^S) \geq 0.23$ for all times past 1984. Before 1984, $\rho_\tau^2(\chi^m,\chi_{,\tau}^S)$ is not significantly larger than $\rho_\tau^2(\chi^m,\chi^{PL'})$, and we note generally that power, correlation, and high-frequency spectral content for the polar motion and atmospheric excitation data all increase sometime after 1984. This indicates that the uncertainty of the data is relatively high before 1984.

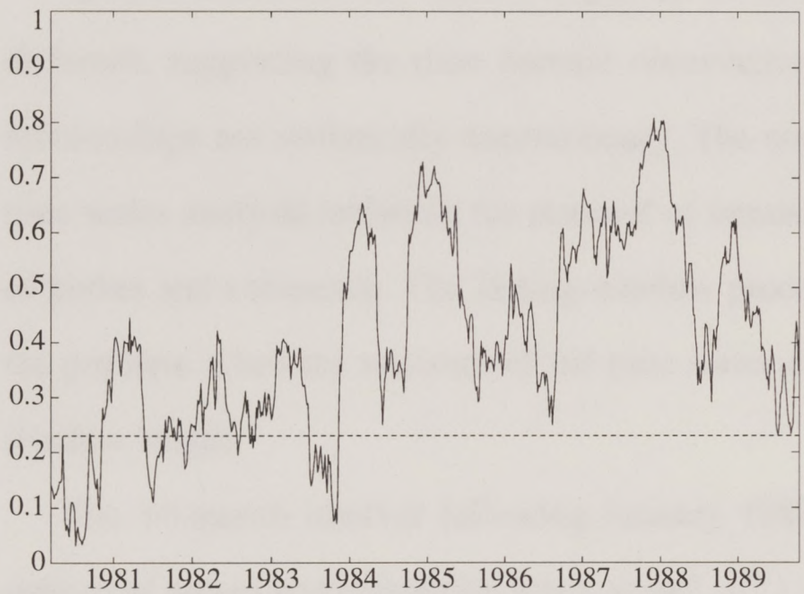


Fig. 7. Complex correlation squared $\rho_\tau^2(\chi^m,\chi_{,\tau}^S)$. $\chi_{t,\tau}^S$ is a linear combination of atmospheric time series for land and ocean regions which minimizes the sum of the squared deviations from χ_t^m for each 150-day window centered on time τ.

Because of the yearly variation of $\sigma_\tau^2(\chi^m)$ and $\sigma_\tau^2(\chi^{PL})$, the presence of noise in either time series may produce a yearly variation in $\rho_\tau^2(\chi^m,\chi^{PL})$. We tested this hypothesis by adding Gaussian white noise to χ_t^{PL} and computing ρ_τ^2 between the original and corrupted time series. The addition of noise with an amplitude of several milliarc seconds reproduces part of the pattern of yearly variation in $\rho_\tau^2(\chi^m,\chi^{PL})$. However, such a noise amplitude in the atmospheric data implies large uncertainty in the planetary-scale atmospheric geopotential heights or, equivalently, spatially correlated noise. Because the noise characteristics of the atmospheric data are

uncertain and the noise level of the polar motion data is well established from independent geodetic methods, we conclude that part of the yearly variation of $\rho_\tau^2(\chi^m,\chi^{PL})$ could be attributed to noise in the atmospheric data. However, some of the yearly variation is real because $\rho_\tau^2(\chi^m,\chi_{,\tau}^S)$ is statistically significant (i.e., at or above the 99% confidence level) past 1984.

Frequency Domain Estimates

Spectral estimates for various segments of the time series are all considerably different, supporting the time domain observations that the time series and their relationships are statistically nonstationary. The nonstationary behavior complicates time series analysis including the removal of trends and sinusoids and the estimation of power and coherence. The sliding-window procedure is a sensible way to handle the problem when the statistics of the time series change slowly with respect to the window length.

The 14-month interval following January 1987 is a special case of the yearly pattern of power and correlation when $\rho_\tau^2(\chi^m,\chi^{PL})$ during the summer is higher than average. The persistence of statistically significant windowed correlation suggests that the data are sufficiently stationary for spectral analysis of coherence and phase over the entire interval. This is the only period in the time series, greater than a few months duration, that is arguably suitable for spectral analysis. The graph of χ_t^m and $\chi_t^{PL'}$ time series is unequivocal evidence that they are correlated (Figure 8). The correlation coefficient between $\mathrm{Re}(\chi_t^{PL'})$ and $\mathrm{Re}(\chi_t^m)$ is 0.67, and for the imaginary part it is 0.83. The magnitude of the complex coefficient is 0.78. The coherence spectrum suggests that about 60-80% of the polar motion and atmospheric excitation

power is linearly related in the range -30 to +12 cpy including the Chandler frequency of 0.843 cpy (Figures 9 and 10). At prograde frequencies between +12 and +24 cpy the time series are essentially incoherent. The phase spectrum is close to zero and generally varies little at frequencies where the coherence is significant, which further supports the statistical significance of the coherence estimates.

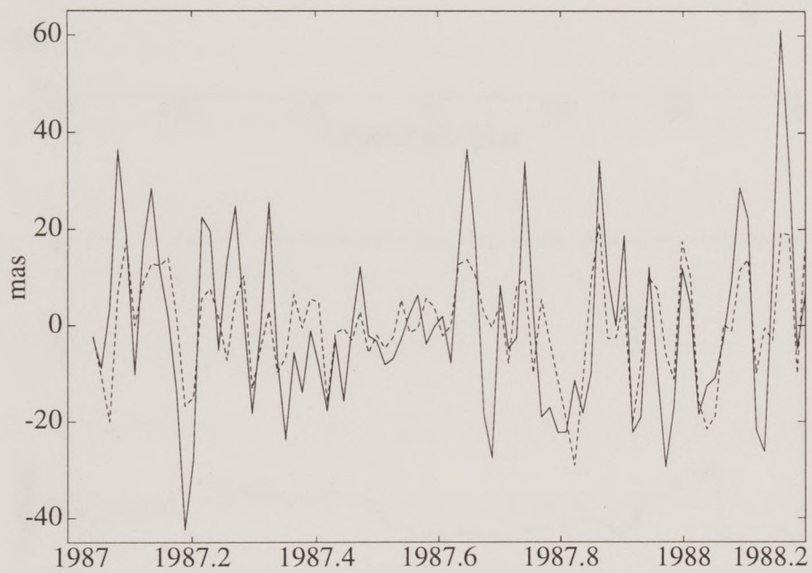


Fig. 8a. Re($\chi_t^{PL'}$) (dashed) and Re(χ_t^m) (solid). The correlation coefficient is 0.67, but $\chi_t^{PL'}$ has about 1/3 the variance of χ_t^n. Spectral analysis of this portion of the time series is shown in Figures 9 and 10.

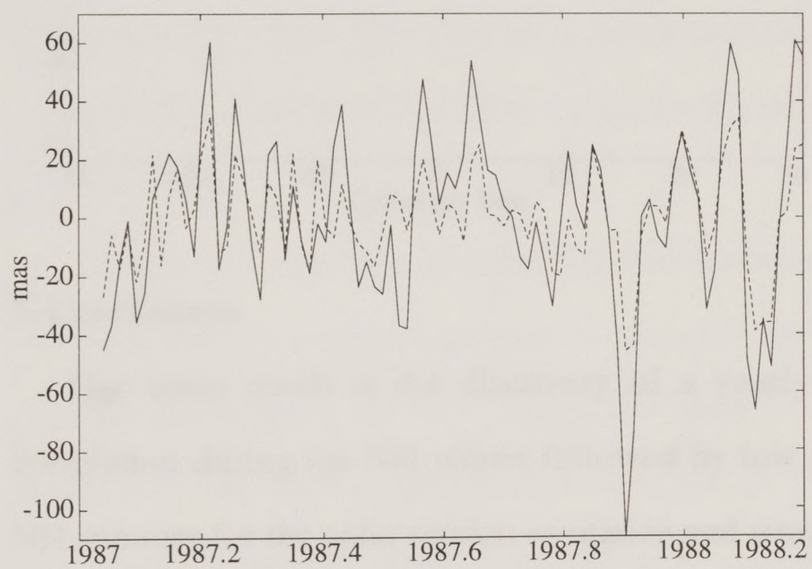


Fig. 8b. Im($\chi_t^{PL'}$) (dashed) and Im(χ_t^m) (solid). The correlation coefficient is 0.84, but $\chi_t^{PL'}$ has about 1/3 the variance of χ_t^m. Spectral analysis of this portion of the time series is shown in Figures 9 and 10.

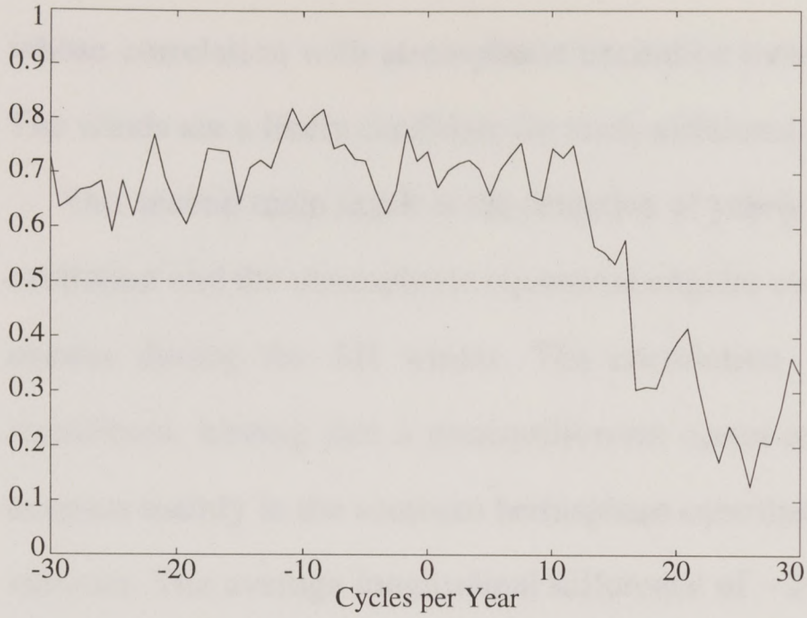


Fig 9. The squared coherence of $\chi_t^{PL'}$ and χ_t^m for January 14 1987 to March 19 1988. The 99% confidence level is 0.5.

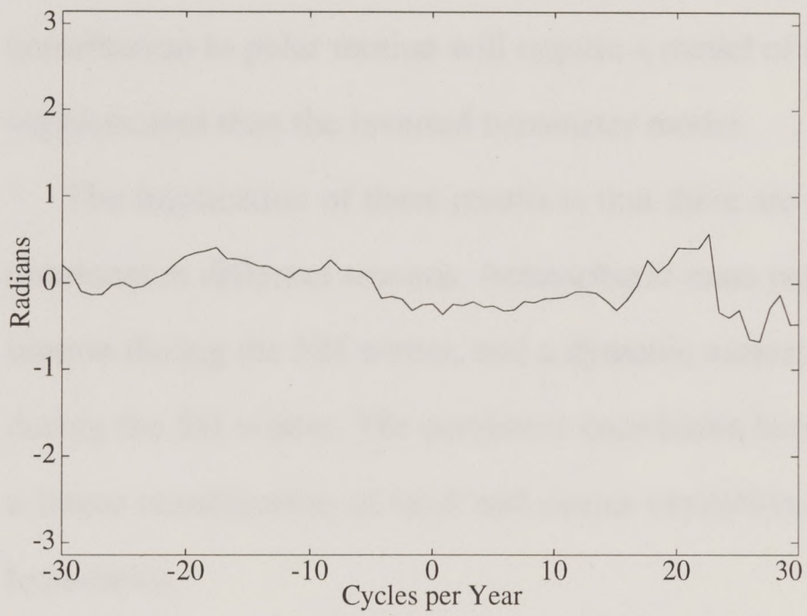


Fig. 10. The phase spectrum of $\chi_t^{PL'}$ relative to χ_t^m. A positive prograde (or negative retrograde) phase corresponds to a lag of χ_t^n with respect to $\chi_t^{PL'}$.

5. Conclusions

Our main result is the discovery of a yearly recurrence of high power and correlation during the NH winter followed by low power and correlation during the NH summer for the polar motion excitation and atmospheric excitation over land. The similarity between the interannual power fluctuations is also striking. The variations in the correlation and variance ratio (Figure 2) imply that there exist other excitations

whose correlation with atmospheric excitation over land changes from year to year. The winds are a likely candidate for such additional excitation.

Our second main result is the detection of yearly correlation between polar motion excitation and the atmospheric equatorial angular momentum over the midlatitude SH oceans during the SH winter. The correlation, although small, is statistically significant, hinting that a nonequilibrium ocean response, winds, and wind-driven currents mainly in the southern hemisphere contribute to the excitation during the NH summer. The average longitudinal difference of -20° may be related to the dynamic ocean response to pressure variations and winds. Determination of the oceanic contribution to polar motion will require a model of the dynamic ocean response more sophisticated than the inverted barometer model.

The implication of these results is that there are two main excitation sources each dominant at different seasons. Atmospheric mass redistribution over land forces polar motion during the NH winter, and a dynamic atmosphere-ocean response is important during the SH winter. The persistent correlation between polar motion excitation and a linear combination of land and ocean excitations (Figure 7) further supports this hypothesis.

Finally, we observe a 14-month period beginning in January 1987 when 72% of the polar motion variance is linearly related to atmospheric excitation over land at frequencies between -30 to +12 cpy. The general shape of the coherence spectrum, remarkably flat with an abrupt drop at prograde frequencies between +12 and +24 cpy, suggests that the spectral separation of prograde and retrograde frequencies is important and that the incoherent part of the spectrum is a genuine feature. At the Chandler frequency the fraction of linearly related power is about 0.7 (a correlation

coefficient of 0.84). This is an important result because it unequivocally exceeds 99% confidence at the Chandler frequency and adjacent frequencies in a wide band. Thus it is good evidence that the Chandler wobble is meteorologically excited.

Estimates of the Chandler Wobble Frequency and Q

Abstract: The Chandler wobble frequency F and quality factor Q are estimated by least squares using polar motion data in combination with proxy excitations derived from atmospheric data. Monte Carlo tests show that this approach can provide useful improvements over traditional maximum likelihood methods, which use only observed polar motion. With less than a decade of good simultaneous polar motion and atmospheric data there is no new information about Q, but 90% confidence intervals for F are near ± 0.004 cycles per year (cpy), making them comparable to maximum likelihood results based on nearly a century of historical polar motion data. Using our preferred proxy excitation we obtain an estimate of F near 0.831 cpy (a period of 439.5 days), which is significantly different from the maximum likelihood value near 0.843 cpy (433 days). This difference can be understood in light of differences in assumptions underlying the two methods.

1. Introduction

This paper is concerned with the problem of estimating the physical parameters which describe the Earth's resonant Chandler wobble (CW). In a terrestrial reference frame, deviations in the equatorial angular momentum of the Earth $\chi(t) = \chi_1(t) + i\chi_2(t)$ and pole position $m(t) = m_1(t) + im_2(t)$ are approximately related by

$$\chi(t) = \frac{i}{\sigma}\frac{dm(t)}{dt} + m(t) \qquad \sigma \equiv 2\pi F(1 + i/2Q) \quad , \tag{1}$$

as a consequence of the conservation of angular momentum [Gross, 1992]. The subscript 1 refers to position along 0° longitude, and subscript 2 refers to position along 90° East longitude, with the complex plane origin near the figure axis. One of the CW parameters is the resonant frequency F, which has an approximate value of 0.843 cycles per year (cpy) or a period near 433 days (d). F is proportional to the frequency of the Earth's rotation, and is a measure of the non-spherical nature of the Earth and its deformation in response to motion of the pole. The other CW parameter is the dimensionless quality factor Q whose reciprocal is proportional to the dissipation of energy near F. F and Q are whole-Earth properties which measure elastic and anelastic behavior at the lowest frequency normal mode observed for the Earth, and hence they are of general interest in geophysics. *Eubanks* [1993] provides a thorough reference for these and many other topics.

Early values of F were obtained by harmonic analysis of $m(t)$, but modern estimates of F and Q are based on a maximum likelihood (ML) method introduced by *Jeffreys* [1940, 1968] using the assumption that $\chi(t)$ is a random Gaussian process for frequencies near F. After removal of seasonal and long period components from $m(t)$, ML estimates of F and Q are those for which $\chi(t)$ has minimum variance. Thus the ML method produces estimates of F and Q without explicit knowledge of $\chi(t)$.

Beginning with papers by *Barnes et al.* [1983] and *Eubanks et al.* [1988], evidence has accumulated that the excitation $\chi(t)$ is related mainly to atmospheric variability. Polar motion is correlated with barometric pressure changes over a broad range of periods extending from several weeks to more than a year. The physical connection between polar motion and barometric pressure change over land is the

change in atmospheric angular momentum associated with atmospheric mass redistribution. Over the oceans the physical connection may be different because the oceans tend to respond to atmospheric surface loading as an inverted barometer, thereby annulling the effect at the sea floor. In a recent study, *Kuehne et al.* [1993] found that barometric pressure changes over land are highly correlated with polar motion during northern hemisphere (NH) winters. However, this correlation disappears during southern hemisphere (SH) winters, at which time pressure changes over SH oceans become correlated with polar motion. The physical connection between pressure changes over the oceans and polar motion remains uncertain given the inverted barometer assumption. Nevertheless, a linear combination of barometric pressure fluctuations from the two geographical regions - land and SH oceans - is correlated with observed polar motion throughout the year. Thus the linear combination may be useful as a proxy estimate for the full excitation. We examine whether the additional information in this and other proxy estimates of the excitation can improve estimates of F and Q.

Our approach is to solve a discrete version of (1) directly for F and Q via least squares (LS). We begin in Section 2 by describing the polar motion and atmospheric data, and develop proxy estimates for $\chi(t)$ in section 3. In Section 4 we develop the LS estimator for the two CW parameters, with associated confidence intervals. With less than a decade of very high quality data available, our LS method has little new to say about Q, but one estimate of F differs significantly from ML values.

2. Data

Polar Motion Data

We use a subset of a polar motion time series called Space93, described in the annual report of the International Earth Rotation Service for 1993. The complex polar motion time series m_t is used to estimate χ_t^m with the two-point convolutional filter [*Wilson,* 1985]

$$\chi_t^m = \frac{ie^{-i\pi FT}}{\sigma T}\left(m_{t+T/2} - e^{i\sigma T} m_{t-T/2}\right), \tag{2}$$

which is a discrete form of (1). The shift of 1/2 the sample interval (T=1 day) poses no problem in alignment as the atmospheric data are available at 1/2-day intervals. The filter (2) has a slight amplitude error with respect to (1) near the Nyquist Frequency, but the error is negligible for this study.

Atmospheric Data

Polar motion excitation arises from mass redistribution and motion relative to the terrestrial reference frame. $\Delta I(t) = \Delta I_{13}(t) + i\Delta I_{23}(t)$ is the complex deviation in the equatorial products of the Earth's inertia associated with mass redistribution, and $r(t) = r_1(t) + ir_2(t)$ is the complex angular momentum of relative motion, such as the momentum of the winds [*Barnes et al.* 1983]. The excitation is

$$\chi(t) = (1.00\Omega\Delta I(t) + 1.43r(t)) / \Omega(C - A) \tag{3}$$

where $C = 7.04 \times 10^{37}\,\mathrm{kg\,m^2}$ is the principal moment of inertia, A is the moment of inertia about an equatorial axis such that $(C - A) = 0.00333C$, $\Omega = 0.7292115 \times 10^{-4}$ radian per second is the mean angular velocity, and the factors 1.00 and 1.43 account for surface loading and rotational deformation.

Estimation of $\Delta I(t)$ is straightforward for the atmospheric mass contribution using surface pressures. We use data from the European Centre for Medium-range Weather Forecasts (ECMWF), consisting of gridded 1000 and 850 millibar uninitialized geopotential heights with a 1/2-day sample interval for the period 1985-93. At each 2.5 degree grid point the surface pressure s for a particular time t is estimated by exponential interpolation as

$$s = 1000 \exp\left[\ln(850/1000)(h_{1000} - h)/(h_{1000} - h_{850})\right] \tag{4}$$

where h is the grid elevation and h_{1000}, h_{850} are the geopotential heights. Surface topography is supplied by the National Center for Atmospheric Research (NCAR), which also distributes the ECMWF data. Using spherical coordinates to evaluate $\Delta I(t)$, the first term on the right hand side of (3) is

$$\chi_t^P = \frac{1.00R^4}{(C-A)g} \iint s_t(\lambda,\phi) e^{i\lambda} \sin\phi \cos^2\phi \, d\lambda \, d\phi \quad , \tag{5}$$

where ϕ is latitude, λ is longitude (increasing eastward from zero at the Greenwich meridian), and $g = 9.8\,\mathrm{m/s^2}$ is the acceleration of gravity at the surface. Our method of estimating s reproduces almost identically the χ_t^P values provided separately by the ECMWF, but eliminates discontinuities associated with ECMWF model changes over

time. Equation 5 is evaluated by summations, usually over the entire globe. However, in this study we obtain separate integrals (summations) over land (χ_t^{PL}) and over SH ($30^\circ - 60^\circ$ S) oceans (χ_t^{PO}). This region of the oceans was found in *Kuehne et al.* to provide essentially all of the correlation between polar motion and atmospheric pressure variability over the oceans.

3. Calculation of Proxy Excitations

We seek an estimate of the excitation $\chi(t)$ containing all contributions (atmospheric mass and winds, oceanic mass and currents, ground water, etc.) to the terms in (3). In the absence of the necessary data, we construct proxy estimates of $\chi(t)$ based on χ_t^{PL} and χ_t^{PO}.

The relationships among χ_t^{PL}, χ_t^{PO},and χ_t^m are described in *Kuehne et al.* [1993] (Figure 1). Briefly, χ_t^{PL} is approximately equivalent to χ_t^P for the entire globe with an inverted barometer ocean response to pressure variations. The variance of χ_t^{PL} tends to be larger during NH winters than during summers. χ_t^{PL} is well correlated with χ_t^m during NH winters, and essentially uncorrelated during summers. χ_t^{PO} has similar properties approximately 6 months out of phase, but correlation with χ_t^m is generally weaker, and somewhat enigmatic because of the presumed inverted barometer behavior of the oceans. The χ_t^{PO} - χ_t^m correlation arises mainly from three SH regions centered on longitudes 20°, 120°, 270°, all proximal to land areas. However, there is little correlation between χ_t^m and atmospheric pressure fluctuations over the adjacent land areas. The three regions do not contribute equally to the correlation at any given time. Rather, for any given SH winter correlation event, one region tends to dominate.

Figure 2 shows the locations of these three southern ocean regions which dominate the correlation with polar motion.

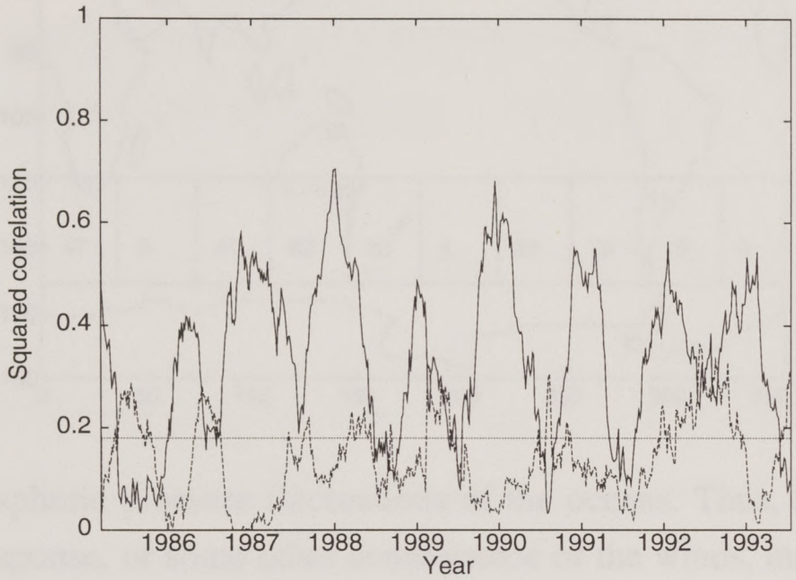


Figure 1a. Correlation coefficients which form the basis for constructing the proxy estimate χ_t^A . The solid line is the squared correlation between χ_t^m and χ_t^{PL}, computed for a 150-day window centered on the date shown. The dashed line is the correlation between χ_t^m and χ_t^{PO}. The correlation for χ_t^{PL} is always high during the winter. χ_t^{PO} follows the same rule 6 months out of phase, although the correlation is generally lower and less predictable. The 99% confidence level is 0.18. The 150-day variance for χ_t^{PL} tends to be high during winter and low during the summer, and *vice versa* for χ_t^{PO}.

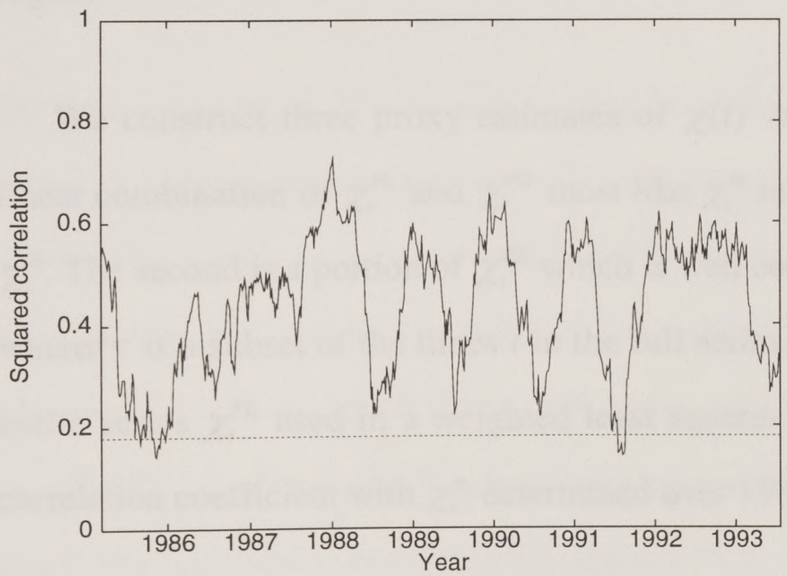


Figure 1b. χ_t^A is a linear combination of χ_t^{PL} and χ_t^{PO} which is correlated with χ_t^m throughout the year.

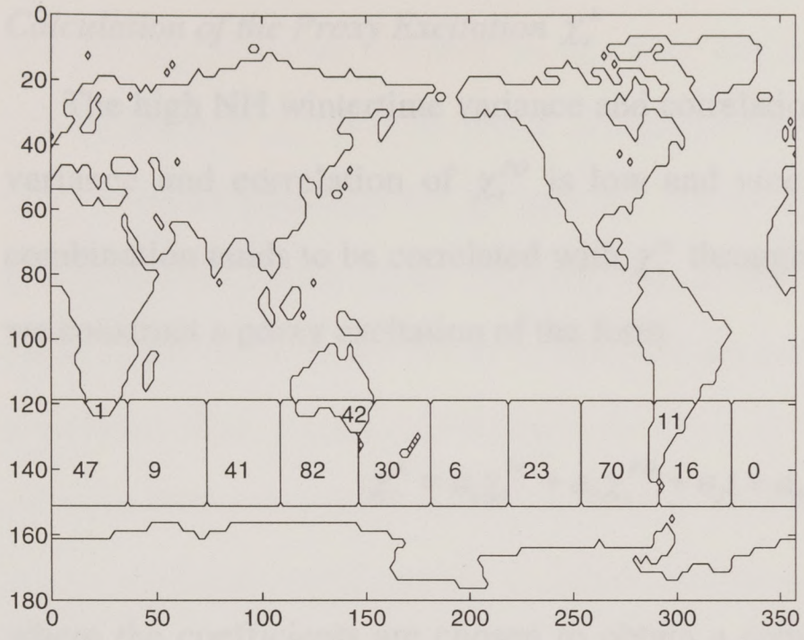


Figure 2. Map of the number of times the squared correlation coefficients exceed the 99% confidence level for the period 1985-93. Correlation is between atmospheric angular momentum (inertia term) over each region and polar motion excitation. The presumed isostatic compensation of the ocean and atmosphere should cancel the effect of atmospheric pressure fluctuations of the oceans. Thus, a dynamic atmosphere-ocean response, or some other consequence of the winds, may be responsible for the correlation. Of all the oceans, only pressure fluctuations over the southern hemisphere region is significantly correlated with polar motion. Although the three most significant regions, centered on longitudes 20°, 120°, 270°, are proximal to land, the atmospheric angular momentum over the nearby land regions is less correlated with polar motion, which suggests that the effect responsible for the correlation occurs over the oceanic region.

We construct three proxy estimates of $\chi(t)$ from χ_t^{PL} and χ_t^{PO}. The first is a linear combination of χ_t^{PL} and χ_t^{PO} most like χ_t^m in a least squares sense, denoted by χ_t^A. The second is a portion of χ_t^{PL} which is well correlated with χ_t^m, denoted by χ_τ^{PL}, where τ is a subset of the times t in the full series. The third proxy excitation is the entire series χ_t^{PL} used in a weighted least squares sense, with weights equal to the correlation coefficient with χ_t^m determined over 150-day windows.

Calculation of the Proxy Excitation χ_t^A

The high NH wintertime variance and correlation of χ_t^{PL} tends to occur when the variance and correlation of χ_t^{PO} is low and *vice versa*. Hence a simple linear combination tends to be correlated with χ_t^m throughout the year (see Figure 1). Thus we construct a proxy excitation of the form

$$\chi_t^A = a_1\chi_t^{PL} + a_2\chi_t^{PO} + a_3 t + a_4 \tag{6}$$

where the coefficients are chosen to obtain a combination of the two atmospheric series which is most like the excitation determined from polar motion via (2). That is, we find the least squares solution to

$$\begin{bmatrix} \chi_{t_0}^{PL} & \chi_{t_0}^{PO} & t_0 & 1 \\ \chi_{t_1}^{PL} & \chi_{t_1}^{PO} & t_1 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} \chi_{t_0}^m \\ \chi_{t_1}^m \\ \vdots \end{bmatrix} \tag{7}$$

for daily time series. The complex valued coefficients a_1 and a_2 serve to scale and rotate (in longitude) the elements of their respective time series, while a_3 and a_4 account for a relative trend and offset. χ_t^m is initially estimated with F=0.843 cpy and Q=175, but this choice is not critical because the series are all broadband. The solution of (7) is found from 8.6 years of data (3195 samples), and gives $a_1 = 0.96e^{8.9^\circ i}$, $a_2 = 0.36e^{-42^\circ i}$, and both a_3 and a_4 nearly zero. Because there are relatively few parameters used to fit the two atmospheric series to χ_t^m, the resulting proxy excitation should contain essentially the same number of degrees of freedom as either χ_t^{PL} or χ_t^{PO}.

The proxy excitation obtained in this way has a broad band squared correlation coefficient with χ_t^m of 0.51, which is significantly higher than the values 0.31 and 0.16 obtained separately for χ_t^{PL} and χ_t^{PO}. In the frequency domain, coherence and phase spectra (Figure 3) are evidence that χ_t^A and χ_t^m are strongly correlated over virtually the entire bandwidth of the data. Retrograde frequencies are somewhat better correlated than prograde, although the high coherence at prograde frequencies between 0.02 and 0.025 cpd (40 to 50 d) is notable. Coherence remains significant across the band containing the Chandler frequency, and the stable phase spectrum further supports the coherence.

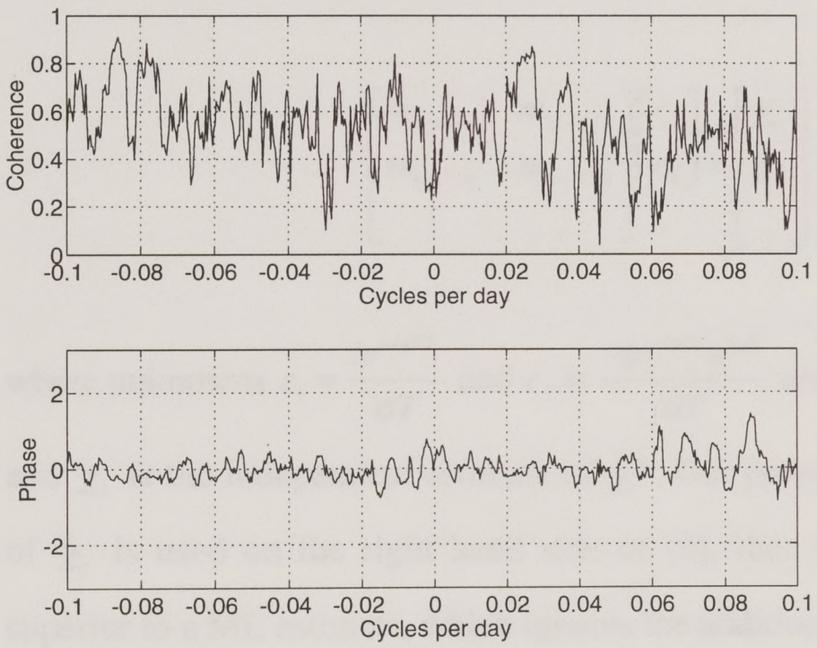


Figure 3. The spectral coherence and phase, computed by a multitaper Fourier method, between χ_t^m and χ_t^A . The statistically significant coherence is broadband, extending from -0.1 to 0.1 cpy, and at the Chandler frequency exceeds the 95% confidence level of 0.4.

Calculation of Proxy Excitations χ_τ^{PL}

The least squares estimator developed in section 4 can be applied to a collection of disconnected portions of χ_t^{PL}. Consequently, for the second proxy estimate χ_τ^{PL} we select portions of the series at those times τ for which χ_t^{PL} is well correlated with χ_t^m. We eliminate those data below a specified cutoff level of squared correlation

coefficients, determined using a 150-day window centered on t as described in *Kuehne et al.* [1993]. We use squared correlation cutoff levels between 0 and 0.34 to form a set of 35 different estimates of χ_τ^{PL}. For example, when the cutoff is 0.018 (the 99% confidence level for the 150-day squared correlation coefficient in the null case) about 20% of χ_t^{PL} is discarded, mainly during NH summers.

4. Estimation of F and Q

The Least Squares Estimator

When independent estimates of χ_t^m and m_t are available, (2) can be written as a set of simultaneous linear equations

$$\begin{bmatrix} m_{t_0+T/2} & m_{t_0-T/2} \\ m_{t_1+T/2} & m_{t_1-T/2} \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \chi_{t_0} \\ \chi_{t_1} \\ \vdots \end{bmatrix} \tag{8}$$

where unknowns $c_1 = \dfrac{ie^{-i\pi FT}}{\sigma T}$ and $c_2 = \dfrac{-ie^{-i\pi FT}e^{i\sigma T}}{\sigma T}$ are the complex coefficients in (2), and χ_t is the independent estimate of χ_t^m. Our proposition is that if a good estimate of χ_t is used on the right hand side of (8), then the solution of (8) ought to be superior to a ML estimate, which ignores the additional information contained in χ_t.

We solve (8) by least squares using one of the proxy estimates of χ_t with additional parameters to account for a residual linear trend, a coordinate system offset, and sinusoidal deficiencies in the proxy at annual and semi-annual frequencies. As before, daily time series are used. The least squares solution is unweighted for the proxy estimates χ_t^A and χ_τ^{PL}, and weighted by 150-day correlation coefficients when

using the entire series χ_t^{PL} as a proxy estimate. LS estimates of the two CW parameters are computed by dividing -c2 by c1 to obtain $e^{i\sigma T}$, i.e.,

$$f = \frac{\arg(-c_2 / c_1)}{2\pi T} \qquad q = \frac{-T\pi f}{\log\left|-c_2 / c_1\right|} \tag{9}$$

where we use the convention that lower case f and q denote estimates of F and Q.

Bias and Confidence Intervals for the LS Estimates

Ensembles of synthetic data can be created using a random number generator in which the true values of the two CW parameters are known precisely, and the LS (or ML) estimator can be applied to the synthetic data to measure estimator performance in terms of bias and confidence interval width. This Monte Carlo approach is useful when knowledge of the excitation is approximate. The Monte Carlo results apply to noise-free data, and thus may underestimate the width of the confidence intervals for real data. However, they provide a valid way to compare the performance of LS and ML estimators.

Before discussing the results of the Monte Carlo evaluation in detail, it is useful to examine the shape of the sum of squares (error-norm) surface as a function of trial estimates of Q and F. A point located at the minimum of this surface is a LS solution to (8). The surface is shown in Figure 4 for a noise-free synthetic data case in which Q =175 and F=0.843 cpy. The error-norm is essentially zero with no prominent minimum for a large range of trial values about the true Q. Thus accurate estimates of Q are difficult to obtain, even with perfect data. Because $1/Q$ is a measure of the Chandler frequency spectral line width, a qualitative explanation for the problem in estimating Q is a familiar one in spectral analysis. A short time series (8.6 years is

only a few Chandler periods) does not contain the information needed to distinguish between a moderately narrow and an extremely narrow spectral peak at the Chandler frequency. In contrast, the shape of the error-norm surface in Figure 4 shows that F is easily determined.

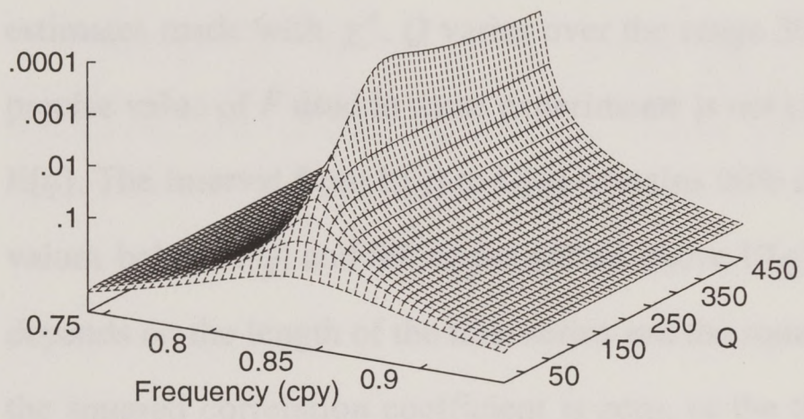


Figure 4. Typical error norm surface of (7) on a $\log_{10}$ scale. Using (2), m_t is deconvolved to χ_t^m for F=0.843 cpy and Q=175. Then m_t is repeatedly deconvolved for a range of trial values of F and Q, and the variance of the difference with χ_t^m is computed. The surface thus graphically describes the behavior of the least squares solution of (7). For trial values of F near 0.843 the error is less than 0.001 for trial Q>100, and on a linear scale the minimum of the surface would not be detected visually. This surface reveals one of the fundamental difficulties of estimating Q using the least squares procedure. If Q is large, then any larger value will do nearly as well in minimizing (7). In contrast F is well determined. These results are closely connected to the Monte Carlo confidence levels in Figure 4.

To evaluate LS bias and confidence intervals for χ_t^A, we first use (3) to transform the 8.6 year series m_t into χ_t^m with known F and Q. We then generate an ensemble of synthetic proxy excitations, χ_t^s, by adding to χ_t^m an ensemble of independent Gaussian noise time series, all having the variance needed to reduce the squared correlation coefficient between χ_t^m and χ_t^s to 0.51 (the same as between χ_t^m and the actual proxy series χ_t^A). From the ensemble of synthetic proxy series, we obtain an

ensemble of estimates f and q via (8). In this process, both positive and negative q are obtained because there is no positivity constraint in (8). Negative q, which are not physically meaningful, occur most often when Q is large and are simply discarded. A similar procedure is used to determine LS bias and confidence for each of the 35 versions of χ_τ^{PL}.

The LS section of table 1 summarizes Monte Carlo statistics applicable to estimates made with χ_t^A. Q varies over the range 30-1000 with F held constant. The precise value of F used in these experiments is not critical. The LS estimate is labeled E[q]. The interval from L[q] to U[q] contains 90% of the ensemble q, with 5% of the values below L[q] and 5% of the values above U[q]. In general, the distribution of q depends on the length of the time series and the squared correlation coefficient. When the squared correlation coefficient is zero, or the time series is short, q tends to be near zero. For the case considered (a squared correlation coefficient of 0.51), q approximately follows a beta distribution biased toward zero (Figure 5). For Q=30 the LS q is about 30, while for Q=100 the LS q is about 65. When Q>100, LS q are hopelessly biased too low. Thus if Q is less than about 70, a reliable LS estimate is possible using χ_t^m and χ_t^A. If Q is much larger, the LS estimator cannot distinguish between the case of large and small Q. In contrast to these problems, the distribution of LS f is normal, unbiased, and reasonably independent of both F and Q, with a 90% confidence interval of ±0.004 cpy for the case where we use χ_t^A.

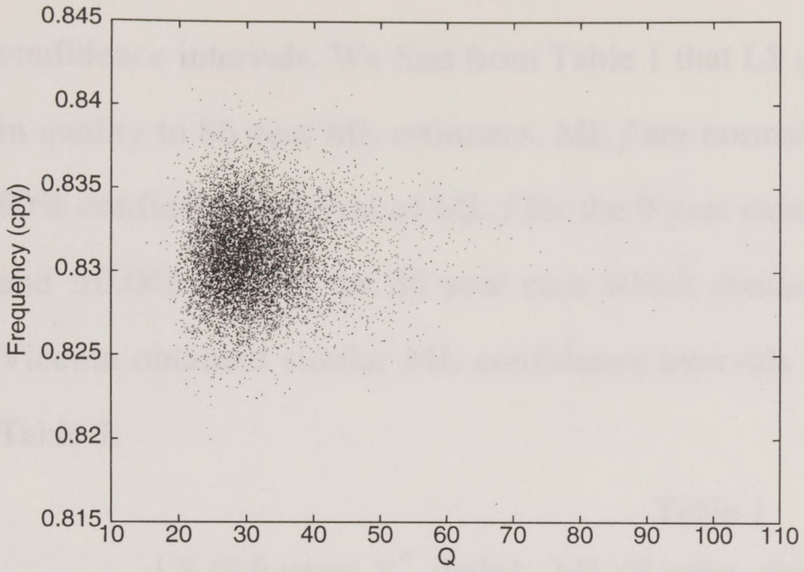


Figure 5. An example of a Monte Carlo experiment described in the text for Q=30. The most likely q is about 28 with a sharp cutoff for smaller q. The probability of obtaining a large q is small, but there is no sharp cutoff. For Q>100 these effects are greatly magnified (see Table 1), and we are left with double jeopardy: the extreme bias means that we will most likely underestimate Q, but neither can we assign a reasonable upper bound.

Using χ_τ^{PL}, 90% confidence for f is about ±0.006 cpy regardless of the correlation cutoff chosen (Figure 6). Thus the confidence gained by omitting uncorrelated data is lost by using fewer data. Similarly, q derived from χ_τ^{PL} are less reliable than for χ_t^A. Although we do not compute them directly, confidence levels for the weighted LS estimate on χ_t^{PL} must be similar to χ_τ^{PL}.

For comparison with the LS estimator, we present Monte Carlo evaluations of Jeffreys' ML estimator described by *Wilson and Vicente* [1990]). The method developed by *Jeffreys* [1968] uses the sequence of Fourier coefficients at 6/7 cpy, computed from non-overlapping 14 month segments of polar motion data, as the basis for the ML estimate. Precisely following the Monte Carlo procedure in Wilson and Vicente, we examine the case of 9 years of daily data, and also the case of 86 years of monthly data, corresponding to estimates made from the historical polar motion data available since the turn of the century. Again, the evaluations are made for the case of noise-free data. The results are summarized in the middle and right sections of Table 1. ML q are biased toward zero for the 9-year case, and cannot be assigned useful

confidence intervals. We find from Table 1 that LS q from 9 years of data are similar in quality to 86 year ML estimates. ML f are normally distributed and unbiased. The 90% confidence interval of ML f for the 9 year case (daily samples) is ±0.004 cpy , and ±0.005 cpy for the 86 year case which contains fewer total data. Wilson and Vicente obtained similar ML confidence intervals for f in their study, as shown in Table 2.

Table 1

	LS (8.6 years χ_t^A, daily) 90% f:±0.004 cpy			ML (9 years, daily) 90% f:±0.004 cpy			ML (86 years, monthly) 90% f:±0.005 cpy		
Q	E[q]	L[q]	U[q]	E[q]	L[q]	U[q]	E[q]	L[q]	U[q]
30	25	25	40	19	16	22	35	20	60
50	45	30	120	22	18	24	50	30	100
70	50	40	230	23	19	25	60	35	140
150	75	55	925	23	20	26	70	40	215
500	85	70	1910	23	20	26	80	45	260
1000	105	75	2160	-	-	-	95	50	270

Table 1. Left section: approximate statistics (nearest multiple of 5) for the LS estimator using χ_t^A (8.6 years of synthetic data), based on 10,000 trials. Middle section: approximate statistics (nearest digit) for the ML estimator for 9 years of synthetic data and 1000 trials. Right section: approximate statistics (nearest multiple of 5) for the ML estimator for 86 years of synthetic data and 1000 trials. E[q] is the most likely value for q, and the interval between L[q] and U[q] contains 90% of the trial values. The LS statistics are somewhat worse for χ_τ^{PL}, which contains fewer data.

Table 2

Estimate	q (90% range)	f (cpy)
LS(χ_t^A, m_t)	72 (30, 500)	0.831 ±0.004
LS(χ_t^{PL}, m_t) (weighted LS)	77 (30,500)	0.836 ±0.006
ML(*Wilson and Vicente* [1990])	179 (74, 789)	0.843 ±0.004
ML(m_t)	-	0.840 ±0.004

Table 2. Confidence for χ_t^A is based on the assumption that χ_t^A contains the Chandler excitation. Confidence for χ_τ^{PL} is larger than ±0.004 cpy because it incorporates fewer data. Confidence for the ML estimates depends on the assumption that the excitation is a Gaussian random process at frequencies near F. Confidence levels for Q using χ_τ^{PL} and the weighted χ_t^{PL} are not computed explicitly, and we state the confidence levels computed for χ_t^A. In both ML and LS cases, Q is poorly determined from 8.6 years of data.

5. Results

Table 2 gives the LS estimates for F and Q obtained using the proxy excitations χ_t^A and χ_t^{PL}, the ML results reported by *Wilson and Vicente*, and the ML results for m_t. Using the proxy estimate χ_t^A, we find an LS value for F of 0.831 cpy with a 90% confidence interval of ±0.004 cpy, corresponding to a Chandler period of about 439.5d. The raw estimate of Q is 52. From Table 1, the bias is about 20, and the corrected value for Q is about 72, with an approximate 90% confidence interval of [30,500]. Confidence is determined from Table 1 using the raw value of 52, which is above the 90% upper limit of Q=30 and below the 90% lower limit of Q=500.

The results for χ_τ^{PL} using a range of squared correlation cutoffs between 0 and 0.35 are shown in Figure 6. f reaches a maximum of .844 cpy near a cutoff of 0.18,

which is the 99% confidence level for squared correlation on a 150-day window. Nearly half of the data are omitted when the cutoff is 0.35, and as this value is approached the least squares solutions deteriorate, producing negatives values for q. Confidence levels for f are constant across the range of cutoffs, implying there is no formal basis for choosing the cutoff. However, cutoffs for χ_τ^{PL} in the 90%-99% range (0.10 to 0.18) provide a reasonable compromise between the desire to omit portions of the atmospheric series which are uncorrelated with polar motion, and the need to retain a sufficient quantity of data . In this cutoff range f is between 0.837 and 0.844 cpy ±0.006 cpy.

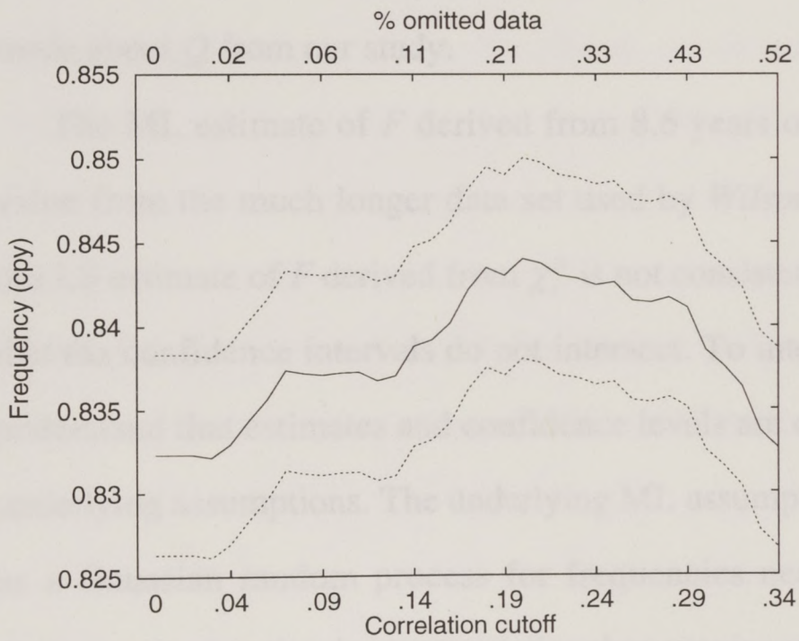


Figure 6. f as a function of cutoff. χ_τ^{PL} contains those portions of data from χ_t^{PL} whose 150-day squared correlation with χ_t^m exceeds the cutoff value. At a cutoff of 0.35 approximately half the data are discarded. Confidence intervals for F increase slightly with cutoff. Bias in estimates of Q is severe for large cutoff.

The weighted LS estimate of F is 0.836 cpy ±0.006 cpy using χ_t^{PL} with a bias-corrected estimate of 77 for Q. This estimate of F is not independent of estimates derived from χ_τ^{PL} because both χ_t^{PL} and χ_τ^{PL} are actually weighted LS estimates. The weights for χ_τ^{PL} are either 1 or 0 (we either accept or reject a datum), whereas every

datum in χ_t^{PL} is used, but with correlation coefficient weights which range between 0 and 1.

The ML estimate of F based on the same 8.6 years of polar motion data used in the LS estimate is 0.840 ±0.004 cpy. The ML Q estimate (19) is meaningless because, according to Table 2, a similar estimate would be obtained for virtually any physically reasonable Q.

In the sense that their confidence intervals overlap, LS estimates of Q are consistent with the ML estimate made by *Wilson and Vicente* [1990] from the longer historical data (Table 2). The quality of the LS estimate, as measured by confidence interval width, is not significantly better. Thus there are no new conclusions to be made about Q from our study.

The ML estimate of F derived from 8.6 years of data is consistent with the ML value from the much longer data set used by *Wilson and Vicente* [1990]. However, our LS estimate of F derived from χ_t^A is not consistent with the ML value in the sense that the confidence intervals do not intersect. To interpret this discordance, one must understand that estimates and confidence levels are conditioned on the validity of the underlying assumptions. The underlying ML assumption is that the excitation behaves as a Gaussian random process for frequencies near the Chandler frequency. The Gaussian assumption is a computational convenience to permit estimation of F and Q without explicit knowledge of the excitation, and it has no basis in observation. The underlying LS assumption is that the proxy excitation is linearly related to the true polar motion excitation, which is strongly supported by the significant broadband coherence and stable phase in Figure 3. Thus, discordance in the estimates may occur, given that the assumptions underlying the estimators are different.

A simple interpretation of the discordance evident in Table 2 is that the ML estimator of F is biased. To understand how this might occur, suppose that the excitation spectrum is not perfectly flat as the ML Gaussian assumption implies, but instead slopes upward between the Chandler frequency and the prograde annual frequency. This slope would cause the spectral peak of polar motion to be slightly asymmetric (skewed toward higher frequencies) with respect to its central value, the resonant Chandler frequency. The ML estimator, which uses a finite-width part of the spectrum to find the center of this Chandler frequency spectral peak, would then yield an estimate which is biased, that is, slightly too high. The LS estimator should have no comparable bias if the proxy for the excitation is a good one. We believe that χ_t^A is a good proxy because of its broadband correlation with polar motion excitation (Figure 3). χ_t^A is preferred over either of χ_t^{PL} and χ_τ^{PL} because it includes a contribution from oceanic regions, although this contribution is not fully understood. Thefore the Chandler frequency of 0.831 cpy (439.5 $\pm$3 mean solar days) obtained from χ_t^A is not only plausible but should be preferable to ML estimates. It is interesting to note that *Jeffreys* [1940] obtained a similar result of 439 $\pm$12 mean solar days.

We conclude that the LS approach can provide useful improvements over traditional ML estimates. That is, there is an important contribution from the additional information in the independent meteorological data used to determine the proxy excitation time series. The LS approach can be applied to recent space-geodetic polar motion data with about the same formal confidence levels for both F and Q as attained by the ML method with 86 years of historical polar motion data. Further improvements in LS results will come with increased time series length, and with

refined estimates of χ_t. Two such refinements may be possible with existing data. The first is to include atmospheric motion effects using wind data. The variable estimates among various meteorological analysis centers suggest that this will not be easy, and we have avoided using wind data for this reason. A second improvement would probably result from better understanding of the physical basis for correlation between χ_t^{PO} and χ_t^m, which implies an improved understanding of the role of the oceans. Beyond these two effects, other data, especially pertaining to fluctuations in water storage among ice, ocean, and land reservoirs, will be required to improve estimates of χ_t.

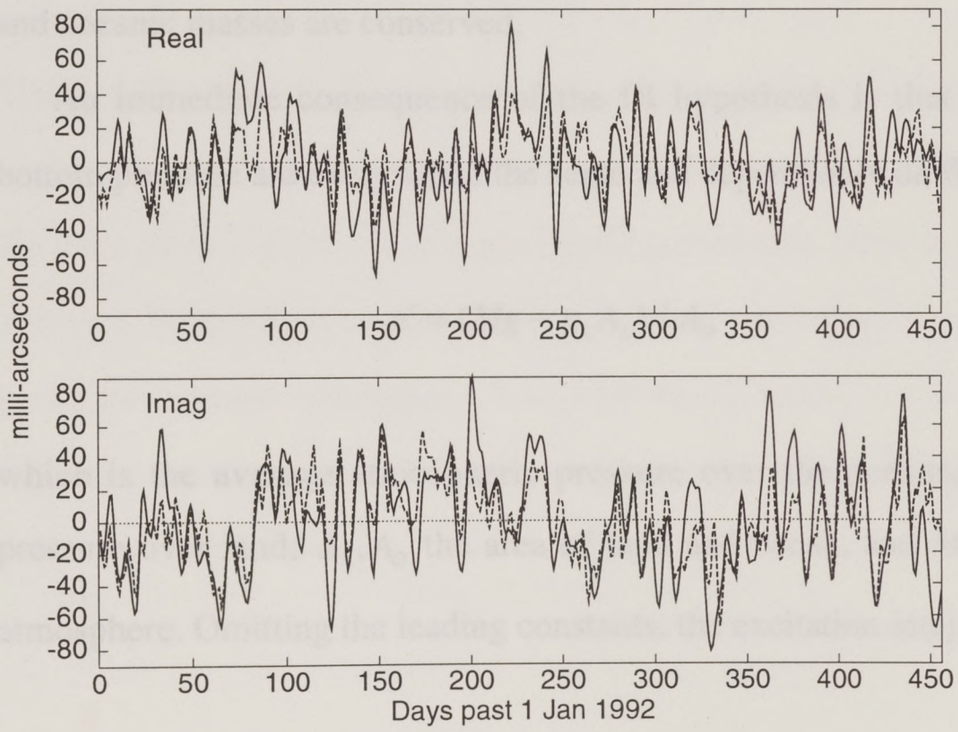


Figure 7. This time domain sample of χ_t^m (solid) and χ_t^A (dashed) is direct evidence that polar motion and atmospheric excitations are closely related.

Appendix A

Calculation of the Inverted Barometer Term

In his pioneering paper, *Jeffreys* [1916] developed a method for computing the polar motion excitation caused by atmospheric pressure fluctuations over the ocean. He assumed that the oceans are in isostatic equilibrium with the atmosphere so that each excess millibar of atmospheric pressure (relative to the mean pressure over the oceans) depresses the accompanying ocean surface ~1 cm. This idea is commonly called the inverted barometer (IB) hypothesis. In this appendix we essentially follow the prescription given by Jeffreys, which for the moment assumes that atmospheric and oceanic masses are conserved.

An immediate consequence of the IB hypothesis is that fluctuations in ocean bottom pressure are everywhere the same and depend only on fluctuations in

$$p' = (Mg - p_L A_L) / A_O \qquad 1$$

which is the average atmospheric pressure over the oceans, with p_L the average pressure over land, A_L, A_O the area of land and ocean, and M the total mass of the atmosphere. Omitting the leading constants, the excitation integral is

$$
\begin{aligned}
&\int_{Earth} p \sin\phi \cos\phi e^{i\lambda} dS \\
&= p' \int_{Ocean} \sin\phi \cos\phi e^{i\lambda} dS + \int_{Land} p \sin\phi \cos\phi e^{i\lambda} dS \\
&= -(Mg - p_l A_l) / A_O \int_{Land} \sin\phi \cos\phi e^{i\lambda} dS \\
&+ \int_{Land} p \sin\phi \cos\phi e^{i\lambda} dS
\end{aligned}
\qquad 2
$$

where the integral on the third line is a consequence of

$$
\int_{Earth} \sin\phi \cos\phi e^{i\lambda} dS = 0 \qquad 3
$$

Thus the IB effect can be computed from atmospheric data over land alone.

One exception to conservation of atmospheric mass is the variable water vapor in the atmosphere. If the water vapor were concentrated entirely over land, the seasonal pressure fluctuations would amount to little more than a millibar, and thus it is probably a small correction. Also, the ocean mass is not constant. As Jeffreys suggests, these problems belong to a comprehensive study of precipitation, runoff, and evapotranspiration.

Appendix B

Computation of atmospheric excitation estimates

Since this research began, many new meteorological data have become available. The old European Centre for Medium-Range Weather Forecasts (ECMWF) initialized geopotential data have been discontinued and are supplanted by the new ECMWF uninitialized data. The new ECMWF surface analysis provides direct estimates of surface pressure (and many other surface parameters) without requiring interpolation from geopotential heights, as described below. The NASA data provide surface pressure estimates 8 times per day on a finer resolution grid. Here is a comparison of angular momentum time series derived from these data which reveals important differences.

Table 1

Agency	Rate	Epoch	Resolution	Data	Symbol
ECMWF	2/day	1980-89	2.5x2.5	Initialized geopotential	χ^{PL}
ECMWF	2/day	1985-93	2.5x2.5	Uninitialized geopotential	χ^{PLG}
ECMWF	2/day	1985-93	2.5x2.5	Surface pressure	χ^{PLS}
NASA	8/day	1985-87	2x2.5	Surface pressure	χ^{PLN}

The atmospheric angular momentum time series in table 1 are evaluated over land regions only. This is nearly the same as the integral over the entire globe using an inverted barometer ocean response, and corresponds to polar motion excitation better than the total integral.

Comparison of χ_t^{PLG} and χ_t^{PLS}

The time series derived from these data do not have the same mean value (Figure 1). Certain systematic differences in the topography data can produce such an offset, such as a slight shift in the land-sea mask that may arise from the ambiguity of defining a coastline at the coarse resolutions used, without strongly affecting the zero-mean part of the time series. For this research the mean value is unimportant, and is subtracted before computing estimates of correlation and phase.

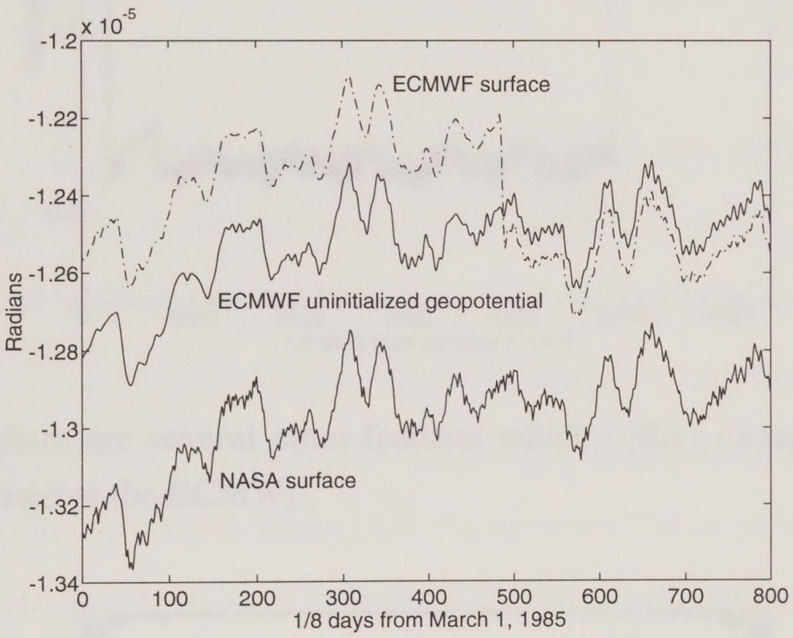


Figure 1. Imaginary (90° east) components of the atmospheric mass angular momentum computed from the ECMWF surface data (χ^{PLS}), the ECMWF 1000 and 850 mbar geopotential heights (χ^{PLG}), and the NASA surface data (χ^{PLN}), for 100 days.

Far more serious are the large jumps and offsets in both real and imaginary components in χ_t^{PLS}, of which one is clearly visible in Figure 1. Several other large offsets are evident when compared with χ_t^{PLG} (Figure 2), or χ_t^{PLN}. These offsets are clearly artifacts of the ECMWF analysis, and are not properties of the atmosphere. Also visible are smaller annual variations in geographic phase and magnitude of the difference.

Aside from the offsets, χ_t^{PLG} and χ_t^{PLS} are nearly identical. The coherence and phase estimates show that the time series are linear versions to within a scale factor

for periods between 4 years and 2 days (Figure 3). At shorter periods the coherence and phase remain highly significant indicators of linear relationship. The power ratio (Figure 4) is essentially unity except at periods less than 2 days.

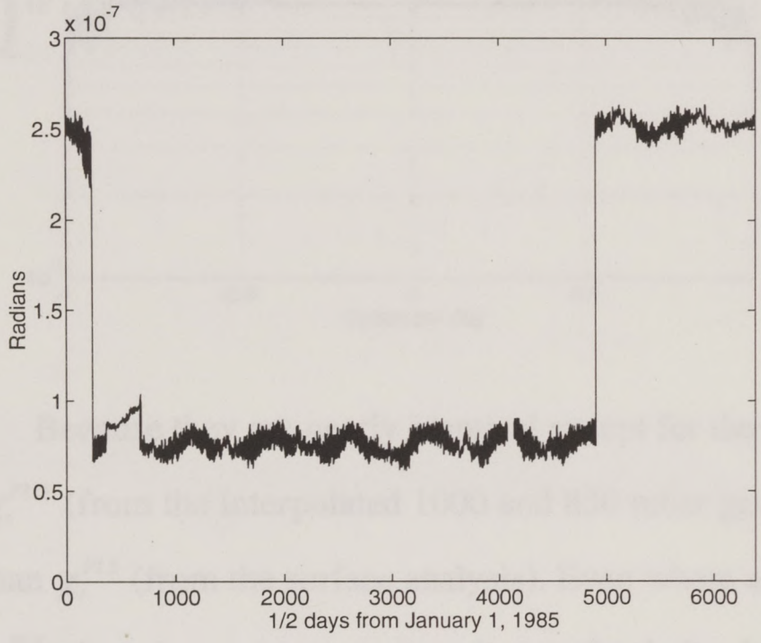


Figure 2. $|\chi^{PLS} - \chi^{PLG}|$. The steps, which are caused by χ^{PLS}, tend to wreck spectral estimates, even though they might not be obvious in plots of the whole time series. There is a small annual variation implying imperfect agreement in the annual sinusoid, which is a large component of both time series. The agreement also fluctuates annually, and there are several other features which reflect changes in the estimation procedures used at the ECMWF.

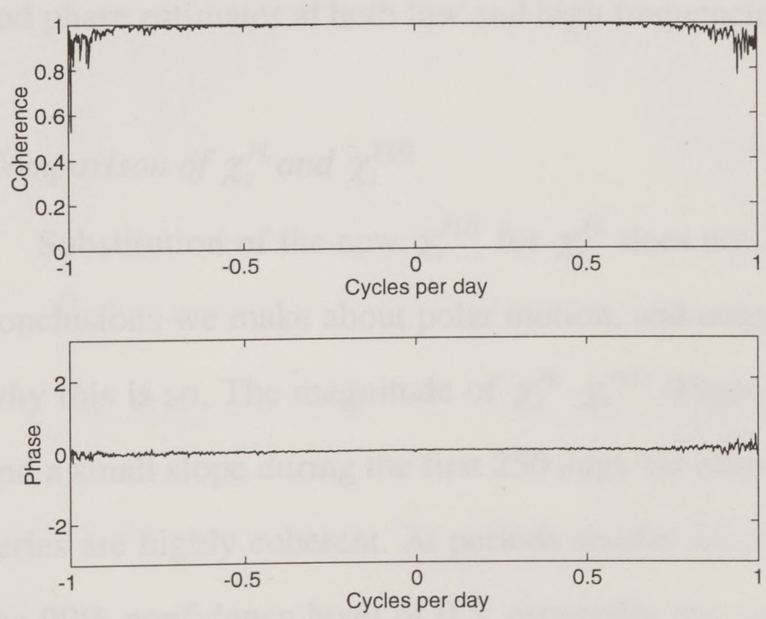


Figure 3. The coherence and phase spectra for χ^{PLS} and χ^{PLG} demonstrate that the time series are essentially linear versions. These estimates are based on a 3000-day segment when χ^{PLS} has no obvious steps, either alone or in relation to the 3 other time series considered.

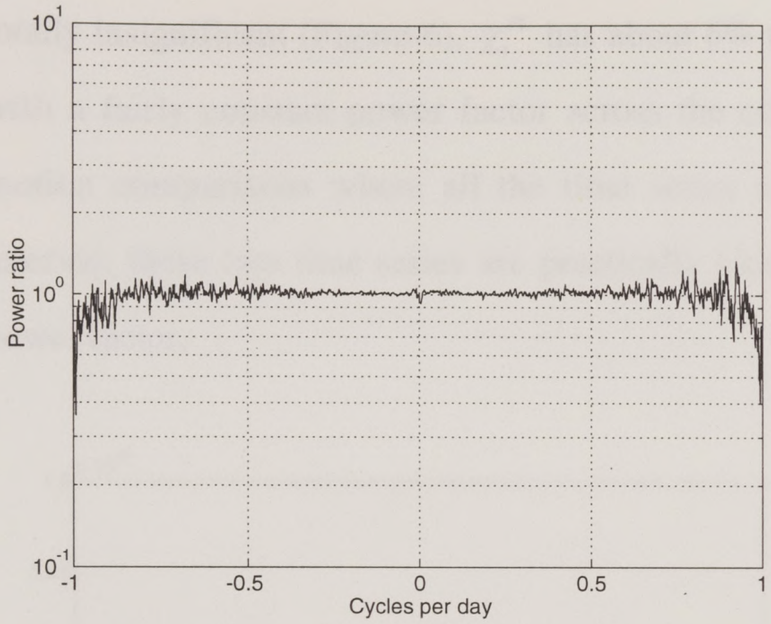


Figure 4. Power spectrum ratio of χ^{PLS} / χ^{PLG}. The ratio is essentially unity up to 0.5 cpd.

Because they are nearly identical except for the errors in the surface pressure data, χ_t^{PLG} (from the interpolated 1000 and 850 mbar geopotential heights) is more reliable than χ_t^{PLS} (from the surface analysis). Even where an offset can be directly detected in χ_t^{PLS}, there is no way to remove it completely by detrending on either side of the step. Because of the noise-like statistics of the time series, a plot of χ_t^{PLS} may not reveal that anything is askew, but the effect of the steps is fairly devastating to coherence and phase estimates at both low and high frequencies.

Comparison of χ_t^{PL} and χ_t^{PLG}

Substitution of the new χ_t^{PLG} for χ_t^{PL} does not significantly change the results or conclusions we make about polar motion, and comparison of these time series shows why this is so. The magnitude of χ_t^{PL}-χ_t^{PLG} (Figure 5) is small, aside from an offset and a small slope during the first 250 days. At periods from 4 days to 5 years the time series are highly coherent. At periods shorter than 2 days the coherence drops below the 99% confidence level of 0.5, especially the prograde part, but it never becomes

totally insignificant (Figure 6). χ_t^{PL} has about 6% more power than χ_t^{PLG} (Figure 7), with a fairly constant power factor across the coherent band. Thus, for the polar motion comparisons where all the time series are decimated to a 5-day sample interval, these two time series are practically identical except for a small constant power factor.

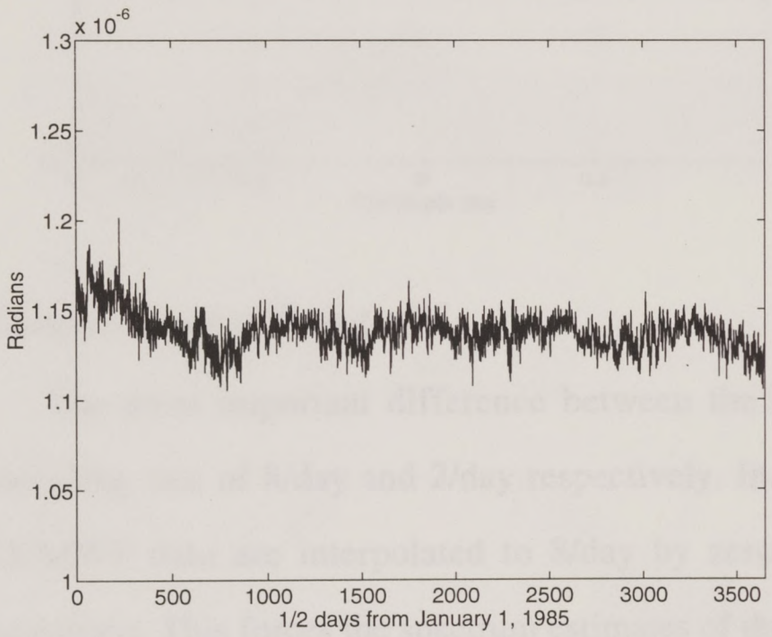


Figure 5. $\left|\chi^{PL} - \chi^{PLG}\right|$.

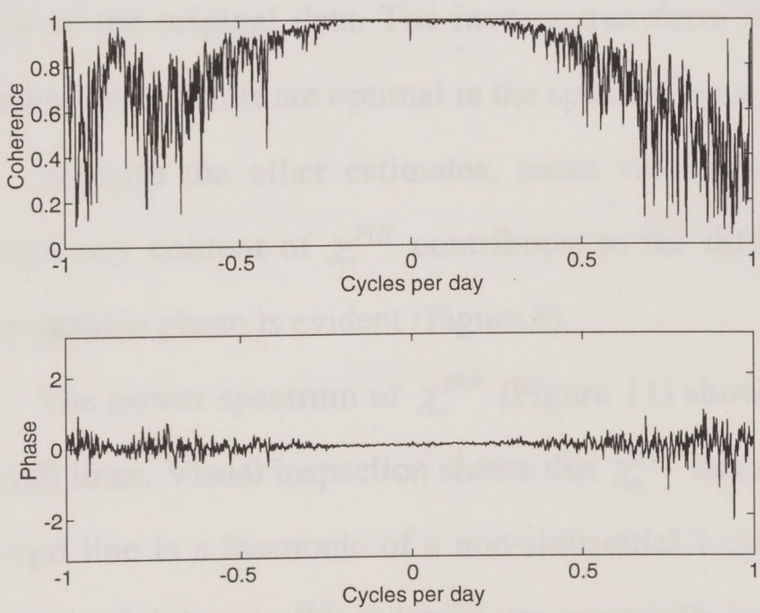


Figure 6. Coherence and phase spectra for χ^{PL} and χ^{PLG}. The times series are essentially linear versions between periods of 8 years and 4 days. The coherence is probably statistically significant at shorter periods. Note the sharp increase in coherence and stability in phase near 0.83 cpd, explained in the text and Figure 11.

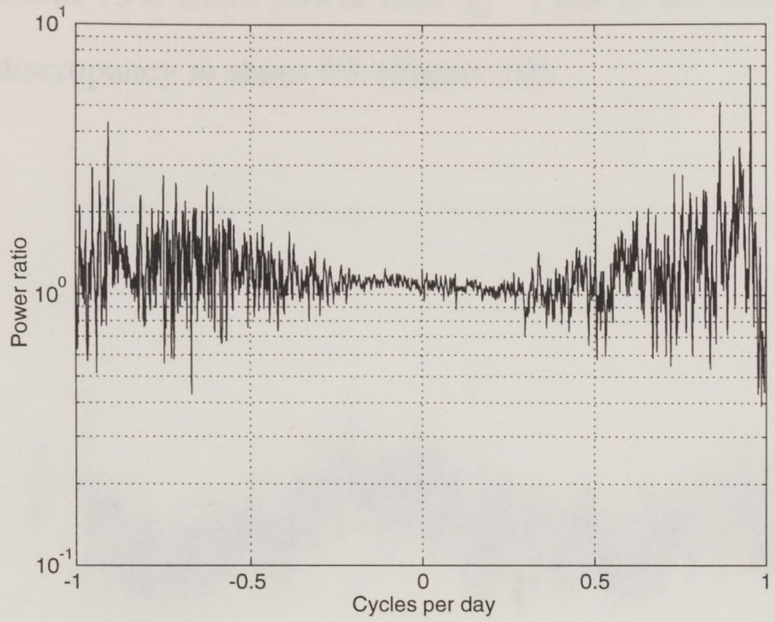


Figure 7. Power spectrum ratio of χ^{PL}/χ^{PLG}.

Comparison of χ_t^{PLN} *and* χ_t^{PLG}

The most important difference between the NASA and ECMWF data is the sampling rate of 8/day and 2/day respectively. In order to compute χ_t^{PLN}-χ_t^{PLG}, the ECMWF data are interpolated to 8/day by zero-padding in the discrete Fourier transform. This forces the spectrum estimates of the interpolated time series to match that of the original data. The inverse transform recovers the original data, and the interpolated points are optimal in the spectral sense.

As with the other estimates, mean values are different (Figure 1). The high frequency content of χ_t^{PLN} contributes to the difference, and a periodic change in geographic phase is evident (Figure 8).

The power spectrum of χ_t^{PLN} (Figure 11) shows strong 1, 2, and 3 cycle per day (cpd) lines. Visual inspection shows that χ_t^{PLN} has a strong 1 and 2 cpd variation. The 3 cpd line is a harmonic of a non-sinusoidal 1 cpd variation. Between periods of 2 years and 4 days, χ_t^{PLN} and χ_t^{PLG} are essentially linear versions (Figure 9). χ_t^{PLN} has

about 15% more power than χ_t^{PLG}, but in the coherent range of -.25 to .25 cpd the discrepancy is about 8% (Figure 10).

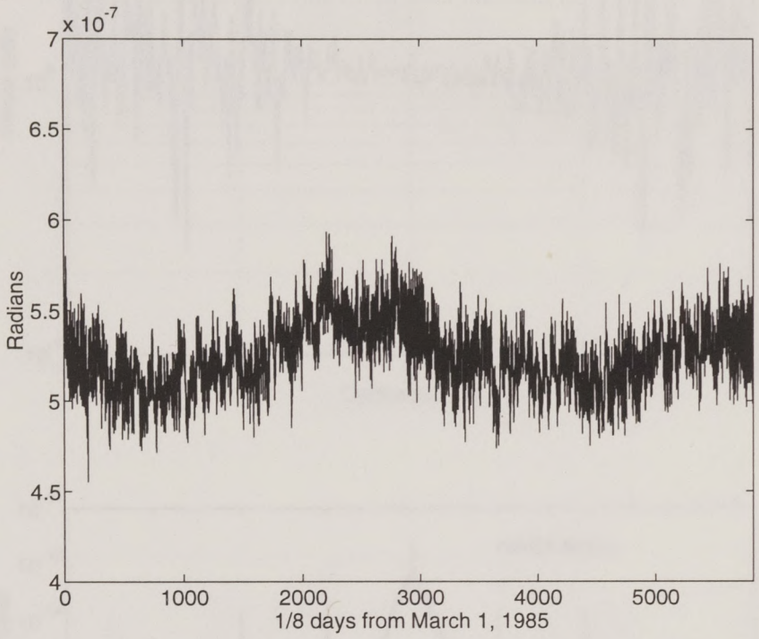


Figure 8. $|\chi^{PLN} - \chi^{PLG}|$.

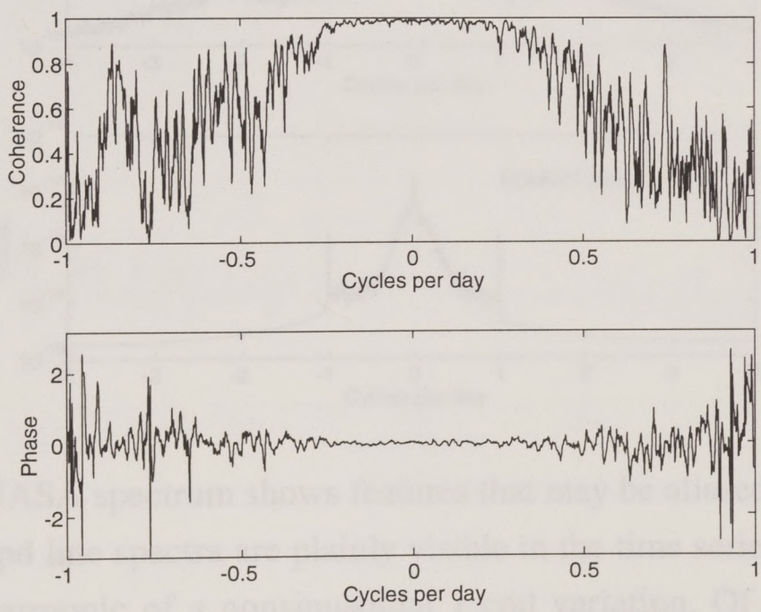


Figure 9. Coherence and phase spectra for χ^{PLN} and χ^{PLG}. The time series are essentially linear versions between periods of 3 years and 4 days. Note the sharp increase in coherence and stability of phase at -0.83 cpd, as explained in the text and Figure 11.

To bring this full circle, Figures 12-14 compare χ_t^{PLN} with χ_t^{PL} (the old initialized ECMWF data). Except for the power ratio, the results are nearly identical to those for χ_t^{PLG}. χ_t^{PLN} has about 8% more power than χ_t^{PL}, but in the coherent range of -.25 to 25 cpd, the power ratio is about 1.

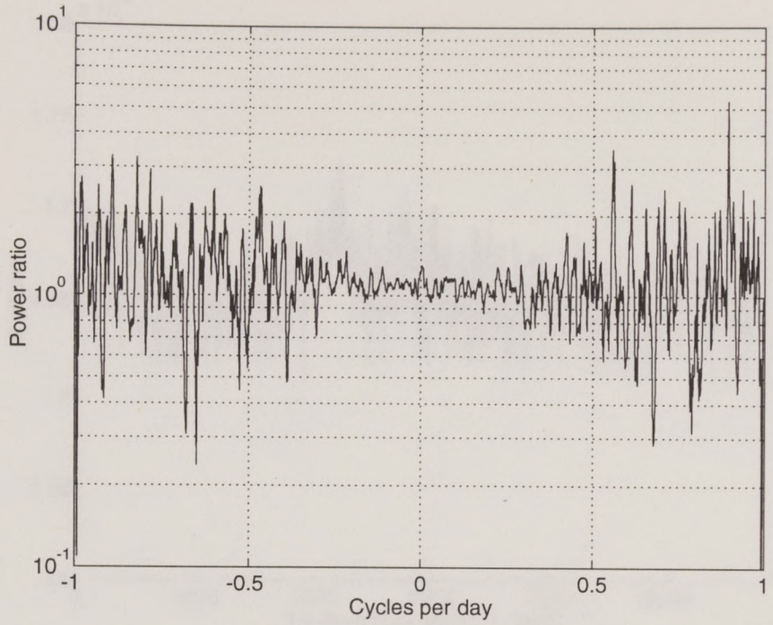


Figure 10. Power spectrum ratio of χ^{PLN}/χ^{PLG}.

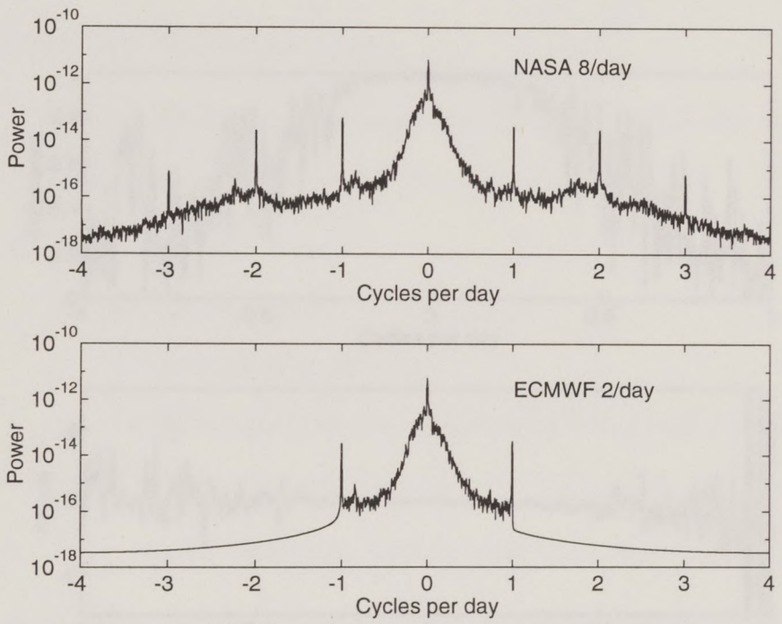


Figure 11. Power spectra for χ^{PLN} (top) and χ^{PLG}. (bottom), both for the 3 year duration of the NASA data available to us. The NASA data are sampled 8 times per day, and the ECMWF twice per day. The ECMWF data are interpolated to 8 times per day to allow a coherence and phase estimate. In the range -1 to 1 cpd the spectra are quite similar. The NASA spectrum shows features that may be aliased in the ECMWF data. The 1 and 2 cpd line spectra are plainly visible in the time series, and the 3 cpd line spectrum is a harmonic of a nonsinusoidal 1 cpd variation. Of special interest is the increase in continuous power near -0.83 cpd in both spectral estimates, with an attending increase in coherence and phase stability (Figure 9). This is one of the retrograde normal modes of the atmosphere which is present in the atmospheric angular momentum because of its harmonic purity.

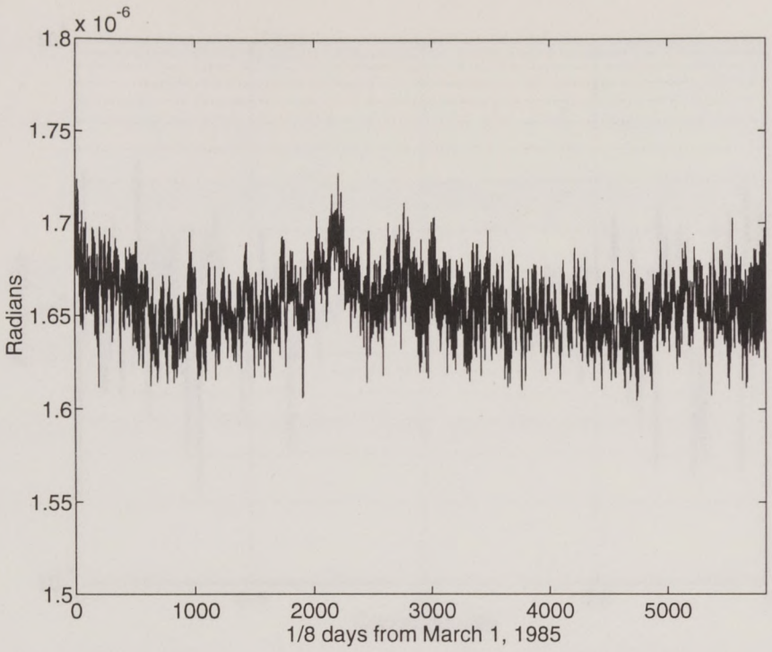


Figure 12. $\left|\chi^{PLN} - \chi^{PL}\right|$.

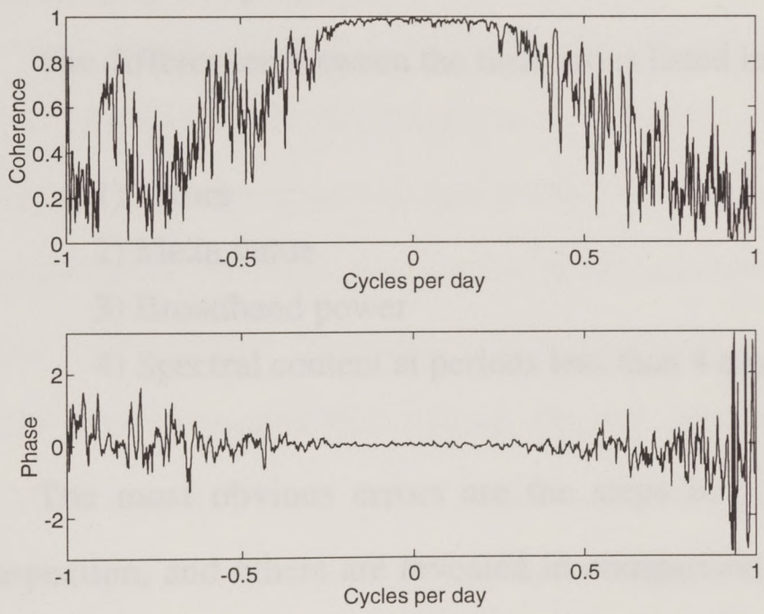


Figure 13. Coherence and phase spectra for χ^{PLN} and χ^{PL}, for the NASA data available to us (1985-87). The time series are linear versions between periods of 3 years and 4 days. Note the increase in coherence and stability of phase near -0.83 cpd.

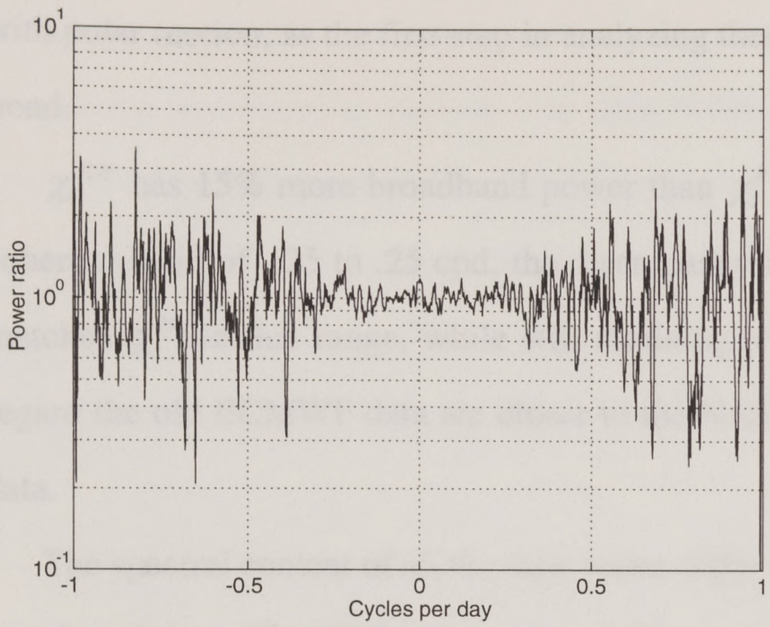


Figure 14. The power spectrum ratio of χ^{PLN}/χ^{PL}.

Summary of Comparisons

The differences between the time series listed in table 1 fall into four classes:

1) Errors
2) Mean value
3) Broadband power
4) Spectral content at periods less than 4 days

The most obvious errors are the steps in χ_t^{PLS}. Some are obvious by direct inspection, and others are revealed in comparison with the other time series. Aside from the steps, this time series is practically identical to χ_t^{PLG}. Because the steps destroy spectral coherence at the lowest and highest frequencies, this time series and their gridded data should be avoided.

All the time series have different mean values. Different treatment of topography data reduction could produce some of the difference, but I can only speculate about the cause of the discrepancies. The mean value itself is unimportant for comparison

with polar motion, as the first step in analyzing these data is removal of the mean and trend.

χ_t^{PLN} has 15% more broadband power than χ_t^{PLG}, and 8% more than χ_t^{PL}. In the coherent band of -.25 to .25 cpd, the discrepancy is reduced about 8%, so that χ_t^{PLN} matches χ_t^{PL} in this range, while still retaining 8% more power than χ_t^{PLG}. In this regard the old ECMWF data are closer to the NASA data the than the new ECMWF data.

The spectral content of all the time series differ at prograde and retrograde periods less than 4 days. The NASA data are sampled at 4 times the rate of the ECMWF data, and show spectral features that may be aliased, or incorrectly estimated in some other way, in the ECMWF time series. One feature that is common to all the power spectra is an increase in continuous power at -0.83 cpd, or a retrograde period of 1.2 days, caused by a barotropic normal mode of the atmosphere (Alquist, 1982). This is evident in all the coherence spectra, which increase to significant levels around -0.83 cpd, with adjacent frequencies at insignificant levels. This is one of several low-wavenumber modes that Alquist did not detect, but which exists theoretically. *Eubanks* [1993] notes that the 1.2d and 10d modes are present in the atmospheric mass angular momentum because they are harmonically pure, and should excite polar motion at periods between 1 and 20 days. In summary, at periods less than 4 days the time series are practically noise except at certain spectral lines and bands corresponding to forced and free oscillations, where the power rises significantly above the continuous background power.

There is good evidence that these time series are essentially linear versions of each other at periods longer than 4 days, up to periods of 2 years (the span of the

NASA data available to us). Because the polar motion data used in this study contain no reliable variations at periods less than 5 days, all the time series except χ_t^{PLS} produce essentially the same results when compared with the polar motion data. Thus the use of χ_t^{PL} in this research is justified, and at the same time the results can be confirmed using χ_t^{PLN} or χ_t^{PLG}.

Appendix C

Neural network system identification

This appendix describes a system identification filter which admits non-linear solutions in a general way, and which requires only minimal information about the functional form of the system. The filter is commonly described as an artificial neural network, but is more closely allied with statistics than with biological entities. The network filter successfully identifies a large class of non-linear systems.

A system is something which takes an input signal and transforms it into an output signal. Audio amplifiers are systems which increase only the amplitude of a signal. Encryption machines are systems which change text into code in such a way that it is difficult to invert the code back into text simply by examining the code. A system is described by a filter, which is a special kind of function that operates on ordered data. A time series, which consists of measurements of some process taken at a uniform time interval, is a natural example of such data.

The task of system identification is to discover a filter that relates observations of the input and output of a system, and validate the results using input-output data that are not used to design the filter. One underlying assumption is that the system does not change. In certain applications, the system identification method should not require a statistically stationary input process. Indeed, we may wish to determine the response of the system to an input process that does not much resemble the input process used to design the filter. On the other hand it is unreasonable to expect that the response of a system to Gaussian white noise can be determined when the design data are sinusoidal. The amplifier is a perfectly linear system which is easy to

identify. A good encryption machine is a highly non-linear system which is difficult to identify.

There are three main classes of applications of system identification: mapping, short-term prediction, and proof of relationship. Mapping is when the identified filter translates the input time series into the output time series with acceptable error at all times. *Narendra and Parthasarathy* [1990] discuss system identification and control in this context. Short-term prediction implies that the filter produces acceptable error for some finite time starting from correct initial conditions, after which time the prediction error increases. *Lapedes and Farber* [1987] and *Tenorio and Wei-Tsih Lee* [1990] investigate prediction of a certain chaotic time series in this connection. Proof of relationship refers to a filter whose error may be too large for mapping or prediction, but the filter significantly boosts spectral coherence for the validation data, thus showing that the input and output time series are related.

This appendix is organized into four sections. The first is a brief introduction to linear system identification. The second is a description of the network filter and a simple method of using it for the system identification problem. The third is a set of demonstrations using specific filters and input data, and the last is the Fortran program used.

Linear System Identification

This appendix is not specifically about linear system identification, but the subject is a good starting point because there is much common ground between linear and network models in system identification. One approach to system identification is to assume that the filter is some linear combination of time-lagged inputs and outputs, i.e.

$$y(t) = c + a_0 x(t) + a_1 x(t-1) + \ldots + b_1 y(t-1) + b_2 y(t-2) + \ldots \qquad 1a$$

where t is an integer time, $x(t)$ is the input and $y(t)$ is the output of the filter. This filter can describe derivatives and integrals, and more generally convolution and feedback, of a time series and hence is quite general. Observations of $x(t)$ and $y(t)$ leads to the matrix equation

$$\begin{bmatrix} 1 & x(1) & x(0) & . & y(0) & y(-1) & . \\ 1 & x(2) & x(1) & . & y(1) & y(0) & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & x(n) & x(n-1) & . & y(n-1) & y(n-2) & . \end{bmatrix} \begin{bmatrix} c \\ a_0 \\ a_1 \\ . \\ b_1 \\ b_2 \\ . \end{bmatrix} = \begin{bmatrix} y(1) \\ y(2) \\ . \\ . \\ . \\ y(n) \end{bmatrix} \qquad 2$$

or $Gm = d$. Practically a least squares solution is sought when the matrix is overdetermined, and there is no linear combination of the columns which produces d. The usual step taken is to find the linear combination of the columns which minimizes the 2-norm magnitude of the difference from d. The solution $\overline{m} = [\overline{c}, \overline{a}_0, \overline{a}_1, \ldots, \overline{b}_1, \overline{b}_2]^T$ is used to model the system as

$$\overline{y}(t) = \overline{c} + \overline{a}_0 x(t) + \overline{a}_1 x(t-1) + \ldots \overline{b}_1 \overline{y}(t-1) + \overline{b}_2 \overline{y}(t-2) + \ldots \qquad 3$$

How well this works depends on many things, including: the linearity of the system; the number of columns forming G; the statistics of the input driving the

system; and noise in $x(t)$ and $y(t)$. For example, some non-linear systems are well described by a linear approximation when the input data are uniformly distributed noise in $\{0,1\}$, but are quite non-linear for input data that are zero-mean Gaussian white noise. Others behave quite linearly when the input is highly correlated, e.g. a sinusoid. If $Gm = d$ is an appropriate model for the system, the number of time lagged terms of $x(t)$ and $y(t)$ can be determined experimentally. Noise, either measurement error or some other unwanted component, is always present in experimental data. As written above, the model treats noise and signal on an equal basis. Finally, the coefficients must be chosen so that the filter output does not grow uncontrollably.

The coherence spectrum of the system input and output is a useful way of quickly assessing whether a linear model of (1b) will succeed in system identification. The coherence is a measure of the fraction of linearly related variance as a function of frequency. If the coherence spectrum is below the 99% confidence level across a broad range of frequencies, linear system identification will probably not succeed at reproducing the behavior of the system at those frequencies. Section 3 describes the method of computing spectral estimates used in this paper, and how to interpret them.

Another useful idea is the response of a system to an impulse. The impulse response for a linear system completely describes the behavior of the system to any input signal. This is not true for non-linear filters. A non-linear system that has an impulse response that rings forever in a complicated way is hard to identify.

In principle non-linear relationships, e.g. $x(t)y(t-1)$, can be included in the linear least squares formulation. This works well for functions which are primarily linear combinations of such terms. The linear method is fundamentally a parameter

estimation method for a known filter form, but for general system identification it is the filter function which requires estimation.

Network filter

This section describes a function which can approximate a large class of non-linear functions, and thus partly fulfills the requirement of being able to approximate the filter function rather than a set of parameters to a known form. Define the function $F(z_j(t))$ as

$$f\left(\sum_k R_k b_k + \gamma\right) = F\left(z_j(t)\right) = \bar{y}(t) \qquad 3$$

where

$$f(w) = \frac{1}{1+e^{-w}} \qquad 4a$$

$$f\left(\sum_i Q_{ki} a_i + \beta_k\right) = b_k \qquad 4b$$

$$f\left(\sum_j P_{ij} z_j(t) + \alpha_i\right) = a_i \qquad 4c$$

where P, Q, R, and α, β, γ are fixed parameters. The non-linear function $f(w)$ has a sigmoidal shape between 0 and 1 as w varies over the real numbers. As any bounded function can be scaled and shifted into this interval, this is not a restriction. The parameters α, β, and γ determine the shift of the sigmoid along the w axis.

When the independent variable $F(z_j(t))$ takes the form $(x(t), x(t-1), \ldots, \bar{y}(t-1), \bar{y}(t-2), \ldots)$, then $F(z_j(t))$ is a filter.

The original reason for using this form was to model a biological neural network. The idea is that each row of P_{ij} corresponds to a neuron whose input stimulus is $z_j(t)$. $f(w)$ is the activation function for the neuron, and the process continues to the next set of neurons corresponding to the rows of Q_{ki}.

The network model is a poor biological paradigm, but it has some interesting functional properties. F simply defines a surface in j dimensions. It is a non-linear mapping functionally composed of a network of identical elements, whose complexity is scalable by increasing the number of elements. Unlike the linear model, the parameters of each element have no easily identified meaning because they don't correspond in any simple way to $z_j(t)$.

The surface defined by F provides a way of explaining how the network model works in system identification. If the $x(t)$ used to identify the system are representative of the possible inputs to the system, and F can map $z_j(t)$ accurately, then the surface so defined will act like a look-up table or interpolator for subsequent $z_j(t)$. Even if these conditions are met, F may not represent the system at all. For example, the true surface may be smooth - continuous and differentiable - but the $x(t)$ used to construct F may undersample it for subsequent $x(t)$.

For many non-linear systems, there exists an interpolation surface which is a good description of the system. If the system is linear, the best thing is to take something like (2), and use linear least squares methods to find the coefficients. For non-linear systems, there is no reason to linearize, except that it gives the wrong answer a million times faster than exploring non-linear solutions.

System Identification using $F(z_j(t))$

The most simple and least efficient method for determining the network parameters is based on a gradient search of the error

$$\varepsilon = \left(\bar{y}(t) - y(t)\right)^2. \qquad 5$$

Thus the gradient method described is a kind of least squares estimate. Although it is not a true steepest descent algorithm, it works well for the examples presented, allowing for the time required to perform hundreds to thousands of million-floating-point operations. It also can be written in a couple of pages of Fortran or C.

To determine the network parameters, the operational form of $z_j(t)$, which contains filter feedback terms $\bar{y}(t-1), \bar{y}(t-2), \ldots$, is replaced with $(x(t), x(t-1), \ldots y(t-1), y(t-2))$, which uses actual observations of the system output $y(t)$. Thus $z_j(t)$ corresponds to a row of G in (2). For a selected $z_j(t)$ and $y(t)$, the gradients

$$\upsilon\frac{\partial\varepsilon}{\partial P_{ij}}, \upsilon\frac{\partial\varepsilon}{\partial Q_{ki}}, \upsilon\frac{\partial\varepsilon}{\partial R_k}, \upsilon\frac{\partial\varepsilon}{\partial\alpha_i}, \upsilon\frac{\partial\varepsilon}{\partial\beta_k}, \upsilon\frac{\partial\varepsilon}{\partial\gamma} \qquad 6$$

are subtracted from their corresponding estimates of P_{ij}, Q_{ki}, R_k, and α_i, β_k, γ, using a small step υ. The $z_j(t)$ are chosen randomly from the set used to design the filter and the process is continued until ε is acceptably small for all the $z_j(t)$ and $y(t)$ pairs used. The gradient search does not guarantee a solution will be found, even if one exists, nor does it guarantee a solution will be optimal. And as already explained,

a solution may not actually represent the system - the crucial part of system identification is testing F with observations of $x(t)$ and $y(t)$ not used to construct F.

In the following derivations, the gradients are expressed in terms of the observables $z_j(t)$, $y(t)$, and functions of them. Starting with (3), $\bar{y}(t) = f\left(\sum_k R_k b_k + \gamma\right)$, and letting $r = \sum_k R_k b_k + \gamma$

$$\frac{\partial \varepsilon}{\partial R_k} = \frac{\partial \varepsilon}{\partial r}\frac{\partial r}{\partial R_k} = \frac{\partial \varepsilon}{\partial r} b_k = \frac{\partial \varepsilon}{\partial \bar{y}}\frac{\partial \bar{y}}{\partial r} b_k = 2(\bar{y} - y) f'(r) b_k = 2(\bar{y} - y)\bar{y}(1 - \bar{y}) b_k \qquad 7$$

which uses the chain rule, the fact that the network parameters are pairwise independent, and that $f' = f(1 - f)$ for $f = 1/(1 + e^{-w})$. Similarly,

$$\frac{\partial \varepsilon}{\partial \gamma} = 2(\bar{y} - y)\bar{y}(1 - \bar{y}) \qquad 8$$

For (4b), $b_k = f\left(\sum_i Q_{ki} a_i + \beta_k\right)$, and letting $q_k = \sum_i Q_{ki} a_i + \beta_k$,

$$\frac{\partial \varepsilon}{\partial Q_{ki}} = \frac{\partial \varepsilon}{\partial q_k}\frac{\partial q_k}{\partial Q_{ki}} = \frac{\partial \varepsilon}{\partial q_k} a_i = \frac{\partial \varepsilon}{\partial b_k}\frac{\partial b_k}{\partial q_k} a_i = \frac{\partial \varepsilon}{\partial b_k} f'(q_k) a_i = \frac{\partial \varepsilon}{\partial b_k} b_k(1 - b_k) a_i$$

$$= \frac{\partial \varepsilon}{\partial r}\frac{\partial r}{\partial b_k} b_k(1 - b_k) a_i = \frac{\partial \varepsilon}{\partial r} R_k b_k(1 - b_k) a_i \qquad 9$$

and $\partial \varepsilon / \partial r$ was evaluated in (7). Similarly,

$$\frac{\partial \varepsilon}{\partial \beta_k} = \frac{\partial \varepsilon}{\partial b_k} b_k(1 - b_k) \qquad 10$$

Finally, for (4c), $a_i = f\left(\sum_j P_{ij} z_j(t) + \alpha_i\right)$, letting $p_i = \sum_j P_{ij} z_j(t) + \alpha_i$,

$$\frac{\partial \varepsilon}{\partial P_{ij}} = \frac{\partial \varepsilon}{\partial p_i}\frac{\partial p_i}{\partial P_{ij}} = \frac{\partial \varepsilon}{\partial p_i} z_j = \frac{\partial \varepsilon}{\partial a_i}\frac{\partial a_i}{\partial p_i} z_j = \frac{\partial \varepsilon}{\partial a_i} f'(p_i) z_j$$

$$= \frac{\partial \varepsilon}{\partial a_i} a_i(1-a_i) z_j = \left[\sum_k \frac{\partial \varepsilon}{\partial q_k}\frac{\partial q_k}{\partial a_i}\right] a_i(1-a_i) z_j = \left[\sum_k \frac{\partial \varepsilon}{\partial q_k} Q_{ki}\right] a_i(1-a_i) z_j \qquad 11$$

where $\partial \varepsilon / \partial q_k$ was evaluated in (9). Similarly,

$$\frac{\partial \varepsilon}{\partial \alpha_i} = \left[\sum_k \frac{\partial \varepsilon}{\partial q_k} Q_{ki}\right] a_i(1-a_i) \qquad 12$$

The procedure is summarized as follows:

Preliminary: Collect data (i.e. observations of $x(t)$ and $y(t)$). Choose the number of lagged terms of $x(t)$ and $y(t)$ forming $z_j(t)$. This step is the only *a priori* information used to construct the filter. Scale and shift $y(t)$ to the interval (0,1), as this is the range of F. Initialize network parameters to random values.

repeat until done

Choose a $z_j(t)$ and $y(t)$ pair, corresponding to a row of (2).

Compute a_i, b_k, and $\bar{y}(t)$.

Compute $\partial \varepsilon / \partial R_k$, saving $\partial \varepsilon / \partial r$ for (8) and 9.

Compute $\partial \varepsilon / \partial Q_{ki}$, saving $\partial \varepsilon / \partial q_k$ for (10) and 11.

Compute $\partial\varepsilon / \partial P_{ij}$, saving $\partial\varepsilon / \partial p_i$ for (12).

Multiply gradients by υ and subtract from their respective parameters.

end repeat

Examples

In this section are several examples of system identification using synthetic data and systems. Synthetic examples are ideal for studying system identification methods because every aspect of the problem is controlled. The examples, summarized in the next table, range from almost-linear to highly non-linear filters.

To facilitate comparison, the examples are prepared with a standard setup as follows: The filters are designed with 500 input-output samples, and tested with 500 independent samples. Both the design and test data are Gaussian white noise with unit variance. Additionally each filter is tested with a sinusoidal frequency-sweep. The network algorithm uses a gradient size of 0.05 for 5,000,000 iterations, except for the fourth example, which continus to 50,000,000 iterations. All spectral estimates are computed with eight tapers as described in the section on spectral analysis. All network parameters are initialized with random weights.

Selecting the terms of $z_j(t)$, consisting of some subset of $\{x(t), x(t-1), \ldots y(t-1), y(t-2), \ldots\}$ is an important issue which depends on the specific system identification task. Any *a priori* information available may facilitate selection, but in some applications guesswork may be required. For the linear method, not only must these terms be selected, but products and powers of the selected terms may be important if the system is non-linear. Setting-up the network filter is easier because it

is already a generalized non-linear function. Over-parameterizing usually is not detrimental to the solution, but it costs time.

There is no quantitative guide for selecting the number of network parameters and their configuration. My experience is to start with i 5-to-10 times j, and k half of i.

Table 1

Filter	j,i,k	$z_j(t)$
$y(t) = (x(t) - x(t-1))^2$	2,20,10	$x(t), x(t-1)$
$y(t) = .5y(t-1) + .6\cos(y(t-2)) + x(t)^3 + x(t-1)^2$	4,20,10	$x(t), x(t-1), y(t-1), y(t-2)$
$y(t) = x(t)x(t-1) + x(t-1)\cos(\pi y(t-2)x(t))$	4,20,10	$x(t), x(t-1), y(t-2)$
	4,20,10	$x(t), x(t-1), y(t-1), y(t-2)$
	4,30,15	$x(t), x(t-1), y(t-2)$
$y(t) = \cos(x(t)) + \cos(y(t-1))$	2,20,10	$x(t), y(t-1)$
	2,30,15	$x(t), y(t-1)$

The first filter in table 1 describes a system which takes the time derivative of a signal and applies a quadratic transfer function. Despite the linearity of the derivative, the filter is quite non-linear. For unit variance Gaussian noise, the coherence spectrum between $x(t)$ and $y(t)$ is below the 99% confidence level, meaning there is insignificant linear relationship (Figure 1.1).

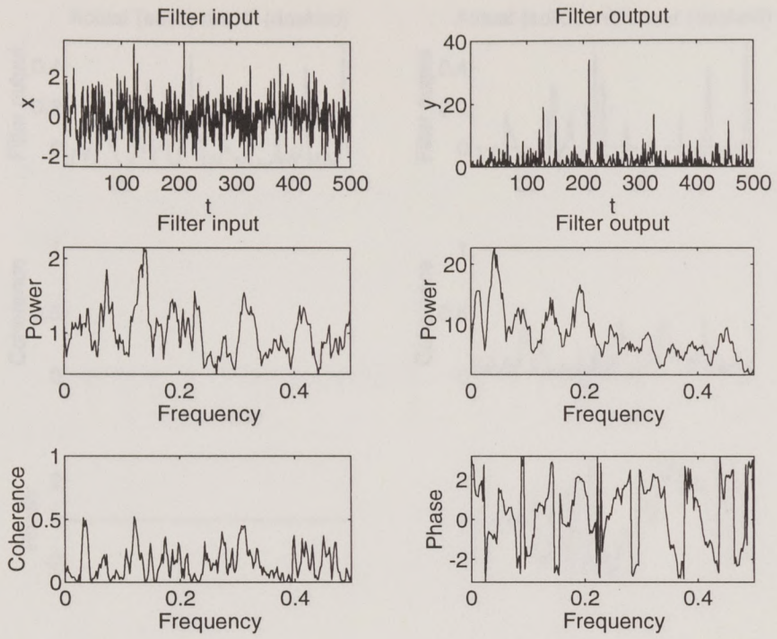


Figure 1.1 **First row**: The filter input is 500 independent Gaussian white noise samples with unit variance. The filter $y(t) = (x(t) - x(t-1))^2$ is the estimated squared time-derivative. **Second row**: the power spectrum of a true Gaussian white noise process has equal power at all frequencies. The relatively small variations in the estimated power spectrum are consequences of the particular instantiation of white noise, and the resolution of the spectral method. **Third row**: the spectral coherence is below the 99% confidence level of 0.5, and the phase apparently varies randomly in its domain of $(-\pi, \pi]$. Hence there is no evidence of linearly related variance between $x(t)$ and $y(t)$ at any frequency.

The linear model $\bar{y}(t) = a_0 x(t) + a_1 x(t-1)$ fails utterly to describe the system, not too surprisingly (Figure 1.2). A linear combination of the quadratic terms $x(t)^2$, $x(t-1)^2$, and $x(t)x(t-1)$ correctly identifies the system, which only shows that the linear method can be used to identify the system when the system is known to within a linear combination.

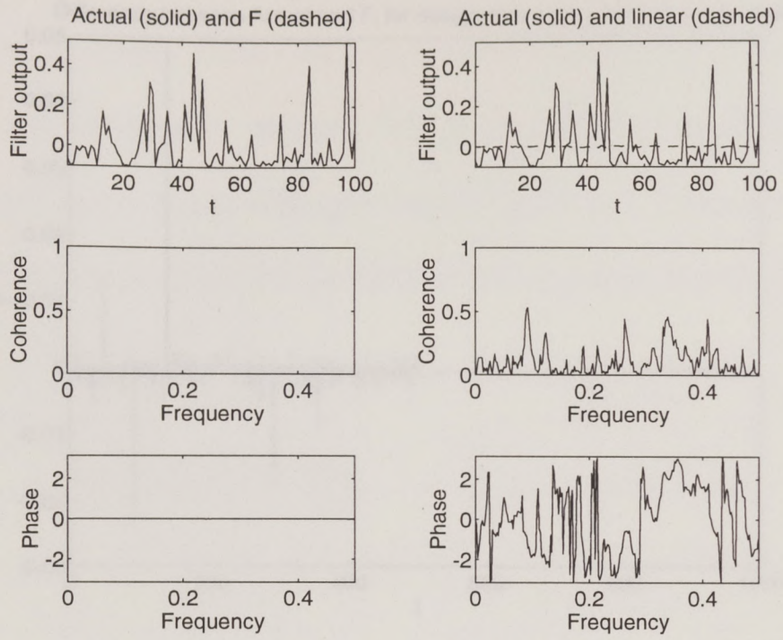


Figure 1.2 **First row**: after designing the network filter F using the $x(t)$ and $y(t)$ data in figure 1.1, the filter is tested with 500 independent Gaussian white noise data, of which the first 100 are shown. Because the network filter has a range of (0,1), the results are displayed in this range, although they can be scaled and shifted back to the proper domain. The network filter is identical to the actual filter. The linear model $\bar{y}(t) = a_0 x(t) + a_1 x(t-1)$ completely fails the test. Explicit inclusion of the quadratic terms is required for the linear method to identify the filter. **Second row**: The coherence spectrum for the network filter F is unity, while that of the linear model remains statistically insignificant. **Third Row**: the filter F has zero phase, and the linear model phase is random, supporting the interpretation of coherence.

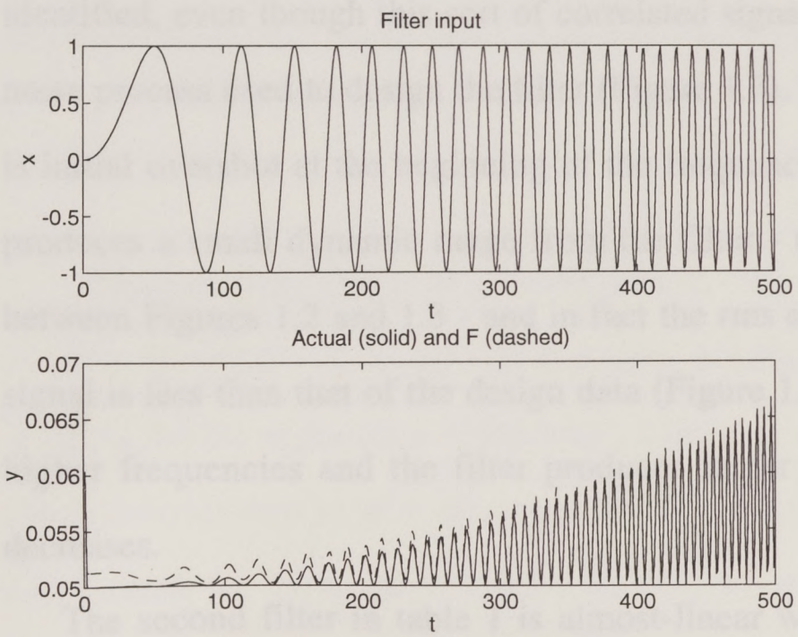


Figure 1.3 The network is tested with a sinusoidal sweep, depicted in the top graph. The actual filter output on the lower graph is well estimated by the network filter, even though this sort of correlated signal is different from the Gaussian noise design data.

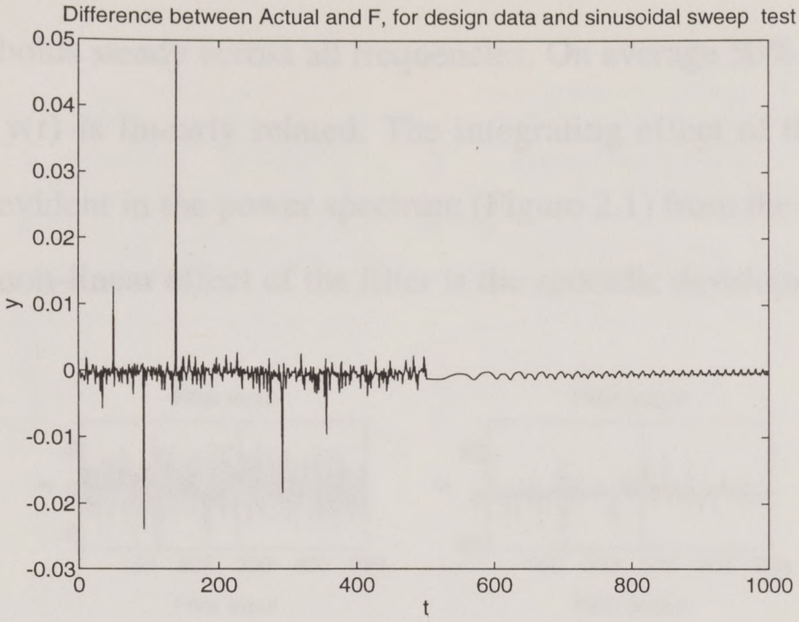


Figure 1.4 The difference between the actual filter and the network filter, for both the design data (0<t<500) and the sinusoidal sweep test. The error is small in both sections for the filter dynamic range of (0,1).

The network filter successfully identifies the system without explicit inclusion of the quadratic terms - only $x(t)$ and $x(t-1)$ are available to the filter input. The rms error between $y(t)$ and the network filter is insignificant, and the coherence and phase are essentially 1 and 0, respectively (Figure 1.2). The variance reduction takes place in the first 100,000 iterations. A frequency-sweep sinusoid input is also correctly identified, even though this sort of correlated signal does not resemble the Gaussian noise process used to design the filter (Figure 1.3). Most of the rms error in this case is initial overshot at the beginning of the frequency sweep. Also the sinusoid input produces a small dynamic range from the filter - note the large difference in scale between Figures 1.2 and 1.3 - and in fact the rms error for the sinusoidal part of the signal is less than that of the design data (Figure 1.4). As the input sinusoid reaches higher frequencies and the filter produces larger output amplitude, the rms error decreases.

The second filter in table 1 is almost-linear when the input time series is unit variance Gaussian noise, based on coherence and phase spectra (Figure 2.1) The variations in the coherence spectrum are probably insignificant because the phase

holds steady across all frequencies. On average 50% of the variance between $x(t)$ and $y(t)$ is linearly related. The integrating effect of the $y(t-1)$ and $y(t-2)$ terms is evident in the power spectrum (Figure 2.1) from the low frequency rise. An important non-linear effect of the filter is the sporadic development of large amplitude spikes.

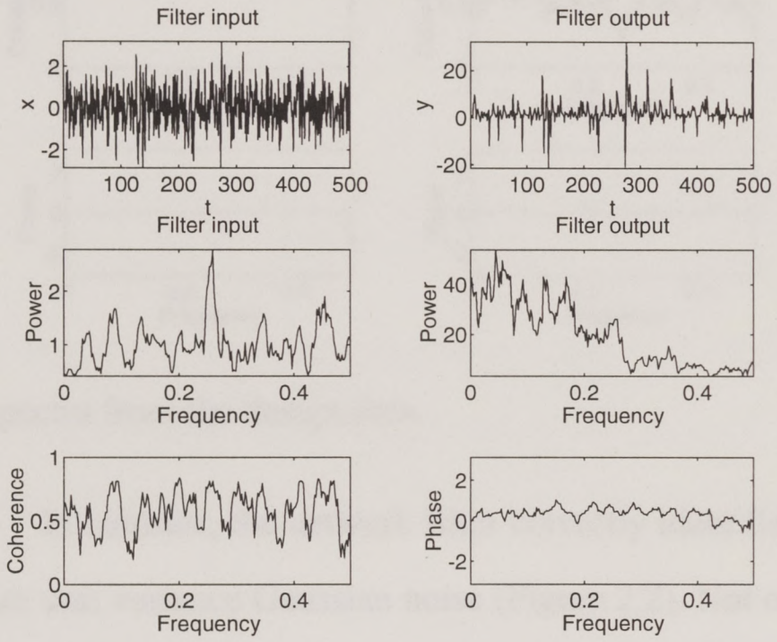


Figure 2.1 The format follows that of figure 1.1, using the filter $y(t) = .5y(t-1) + .6\cos(y(t-2)) + x(t)^3 + x(t-1)^2$. The power spectrum of the filter output clearly shows the integrating effect of the $y(t-1)$ term. The coherence spectrum averages slightly above the 99% confidence level, and the steady phase confirms that these are statistically significant estimates which can be interpreted as fraction of linearly related variance.

Although the filter is almost-linear, the linear model $\bar{y}(t) = c + a_0 x(t) + a_1 x(t-1) + b_1 y(t-1) + b_2 y(t-2)$ is a poor description of the system. The rms error is large, and the application of the linear filter to $x(t)$ does not appreciably change the spectral coherence to $\bar{y}(t)$ (Figure 2.2). The reason is that the linear least squares coefficients for $x(t)$, $x(t-1)$, and $y(t-2)$ are quite small compared to that for $y(t-1)$. $\bar{y}(t)$ is mainly a rescaled version of $x(t)$, because $y(t-1)$ contains most of the linearly related variance.

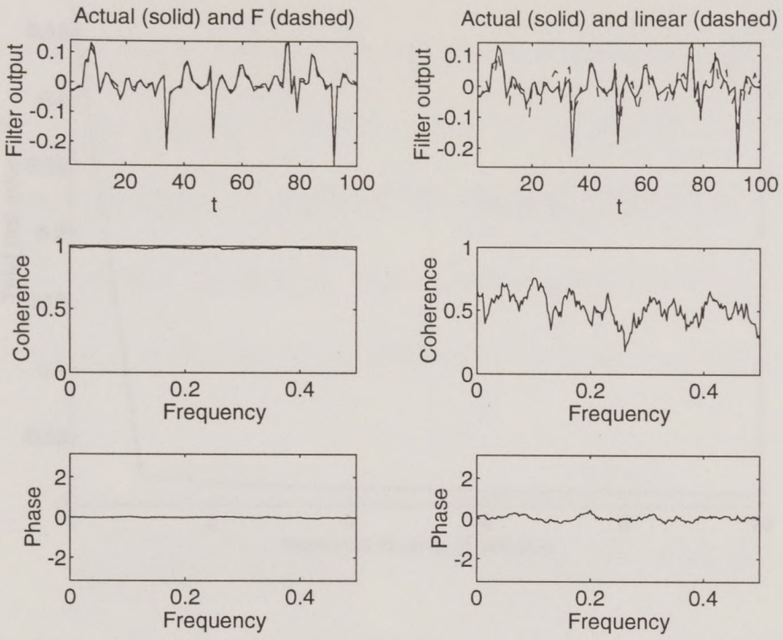


Figure 2.2 The format follows that of figure 1.2. The network filter flawlessly identifies the actual filter. The linear model $\bar{y}(t) = c + a_0 x(t) + a_1 x(t-1) + b_1 y(t-1) + b_2 y(t-2)$ should be a good estimate of the actual filter because of the large fraction of linearly related variance. However, there is little improvement in the coherence and phase spectra from the design data.

In contrast, the network filter correctly identifies the system when the input data are unit variance Gaussian noise (Figure 2.2). Not only is the rms error small, but the spectral coherence and phase are nearly perfect. Starting from an initial random set of network parameters, most of the variance reduction is accomplished in the first 100,000 iterations, or a few minutes on a machine that executes 1,000,000 floating point operations per second. After 1,000,000 iterations there is essentially no further variance reduction (Figure 2.3).

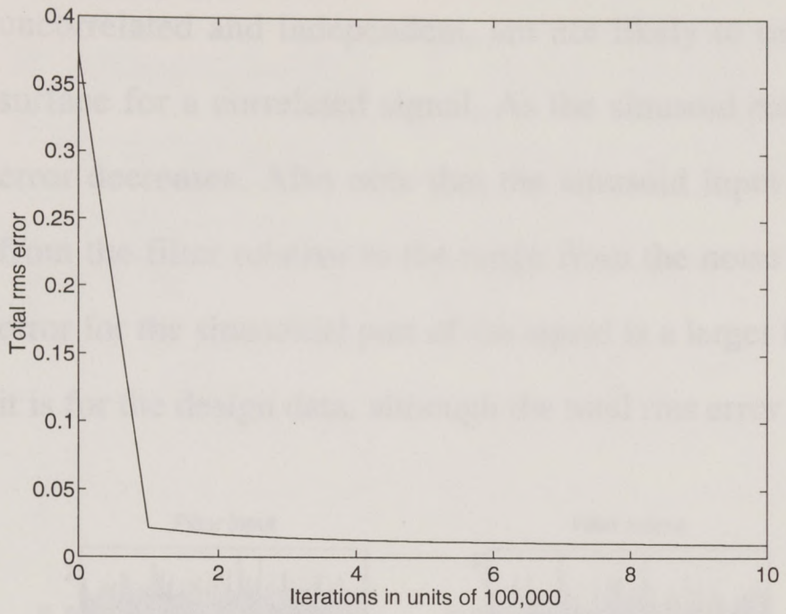


Figure 2.3 The total rms error between the network filter and the actual filter design data reaches nearly its final values during the first 100,000 iterations. As shown in the fourth example, there is no way to decide of the basis of such a graph as to when to cease iteration.

A frequency-sweep sinusoid input is also correctly identified, even though this sort of correlated signal does not resemble the Gaussian noise process used to design the filter (Figure 2.4).

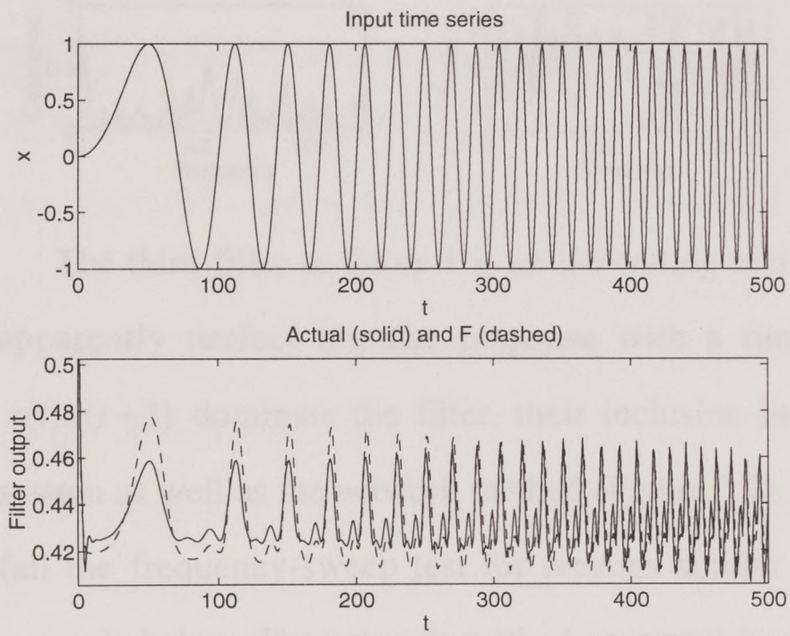


Figure 2.4 The sinusoidal sweep test is well identified by the network filter F, even though this sort of correlated input does not resemble the Gaussian noise design data.

Most of the rms error in this case is overshot along each sinusoid. Qualitatively this error is consistent with the idea of an interpolation surface. The design data, which are

uncorrelated and independent, are not likely to map in any detail the interpolation surface for a correlated signal. As the sinusoid reaches higher frequencies, the rms error decreases. Also note that the sinusoid input produces a small dynamic range from the filter relative to the range from the noise input design data. Hence the rms error for the sinusoidal part of the signal is a larger fraction of the dynamic range than it is for the design data, although the total rms error of each segment is similar.

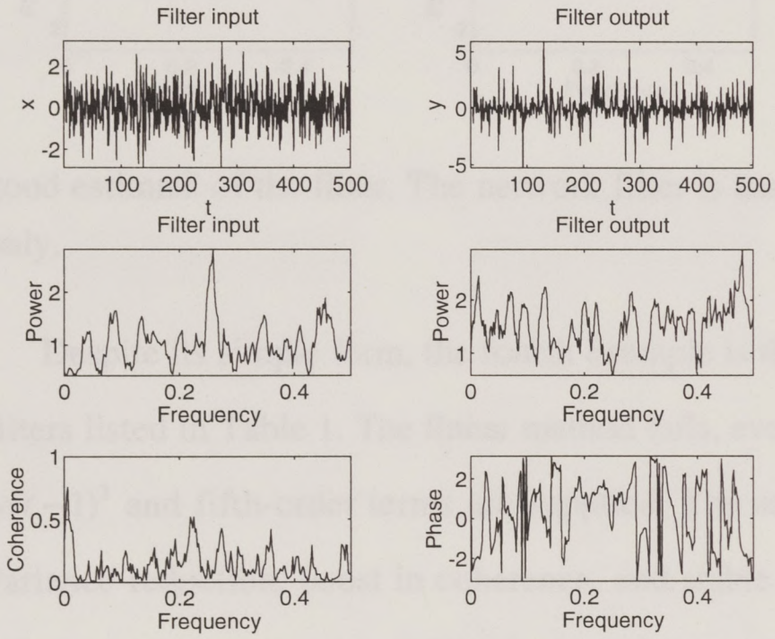


Figure 3.1 The format follows that of figure 1.1. Both $x(t)$ and $y(t) = x(t)x(t-1) + x(t-1)\cos(\pi y(t-2)x(t))$ have fairly white spectra. The time series are linearly related near the frequency of 0.2, based on the stable phase and increased coherence, but outside this band there is no evidence for relationship.

The third filter in Table 1 is an interesting non-linear example because it has an apparently perfect impulse response with a time-lag of 1. Because terms like $x(t)x(t-1)$ dominate the filter, their inclusion in the linear method identifies the system as well as the network method (Figure 3.2). Both linear and network methods fail the frequency-sweep test for reasons similar to those discussed in the fourth example below. The network method accomplishes most of the variance reduction in the first 100,000 steps. Both over-parameterization and a larger network produce identical results. Thus for both linear and network filters, identification only succeeds when the input time series is statistically similar to the Gaussian noise design data.

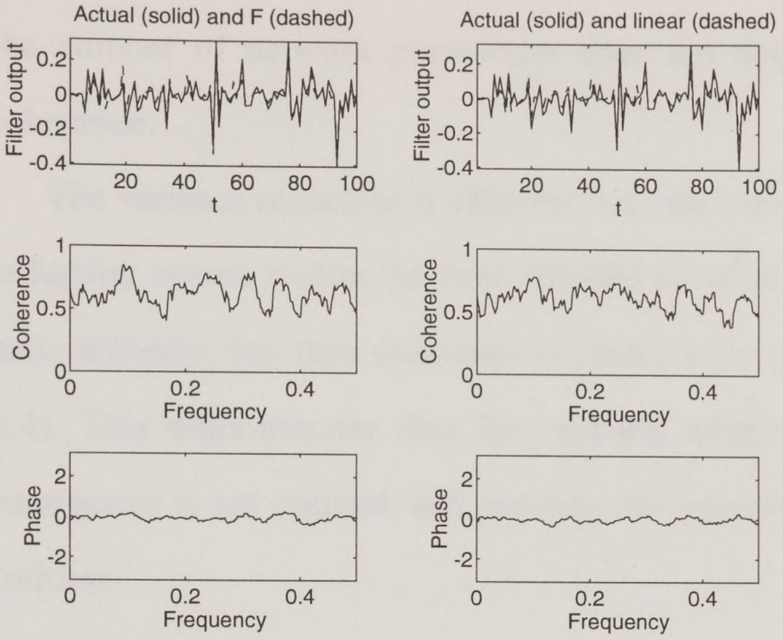


Figure 3.2 The format follows that of figure 1.2. The network filter boost coherence above the 99% confidence level and produces nearly zero phase, as does the linear model which includes $x(t)$, $x(t-1)$, $y(t-2)$, and all possible products of these terms. The product terms dominate the filter, and hence a linear combination of these terms is a good estimate of the filter. The network filter is based on $x(t)$, $x(t-1)$, and $y(t-2)$ only.

Despite its simple form, the fourth example is the most difficult to identify of the filters listed in Table 1. The linear method fails, even when the cubic terms $x(t)^3$ and $y(t-1)^3$ and fifth-order terms are supplied. The network filter produces significant variance reduction, boost in coherence, and stable phase - enough to prove that the two time series are related. The frequency sweep fails for low frequencies, but the agreement becomes better at high frequencies, where the rapid change more closely approximates the Gaussian design data. A plot of the frequency-sweep input and output demonstrates why this filter is difficult to identify (Figure 4.3). The input signal is smooth and correlated, but the output sporadically develops instabilities, characterized by large rapid fluctuations. The corresponding interpolation surface must also fluctuate rapidly over part of its domain. Hence this surface is difficult to map, especially with a Gaussian noise input, because the parts of the map corresponding to smooth changes of input signal are rarely encountered. Increasing

the number of network parameters does not decrease the rms error or increase coherence.

The variance reduction is different from the previous examples, where most of the reduction occurs within the first 100,000 iterations. In this example, the reduction halts abruptly, but then continues suddenly after about 3 million iterations (Figure 4.4). This demonstrates that the gradient method used to estimate the network parameters is not optimal, and provides no indication as to when further iteration is fruitless.

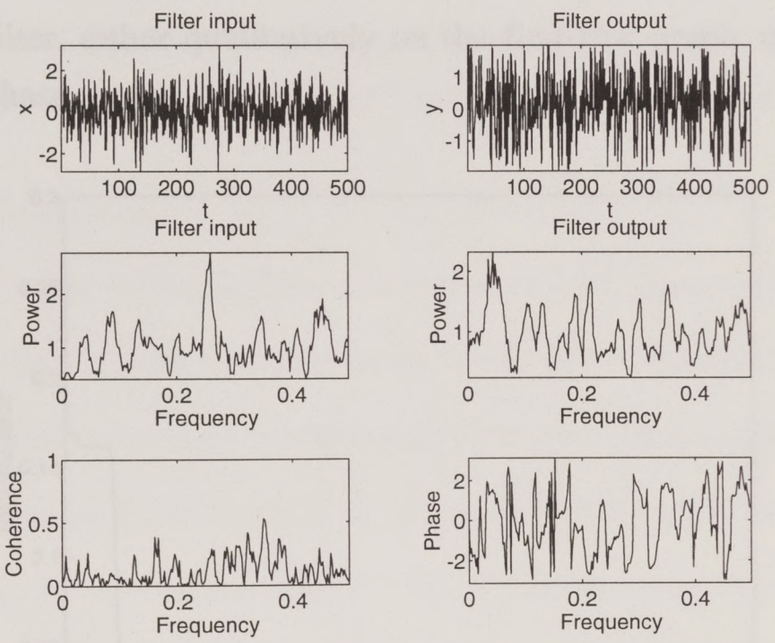


Figure 4.1 The format follows that of figure 1.1. Both $x(t)$ and $y(t) = \cos(x(t)) + \cos(y(t-1))$ have fairly white spectra, and evidently there is no linear relationship between them.

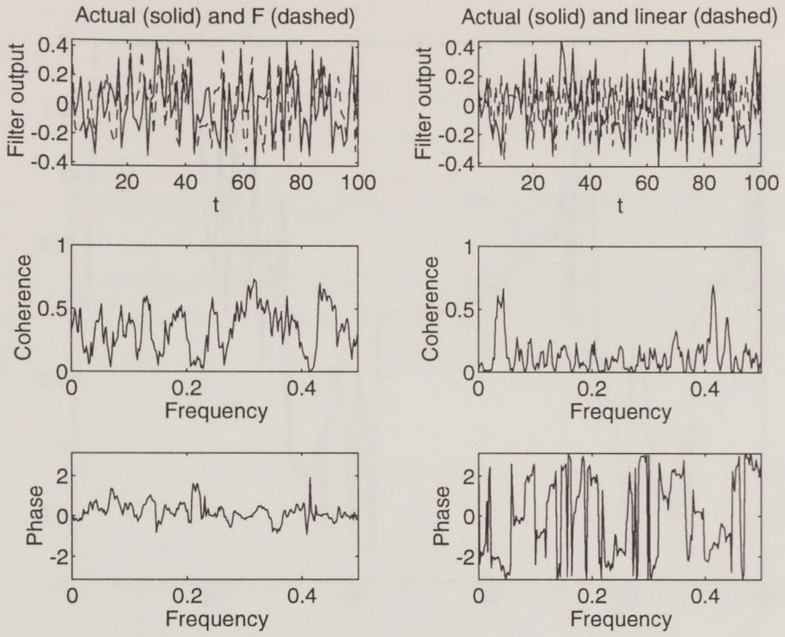


Figure 4.2 The format follows that of figure 1.2. The network filter F resembles the actual filter

$$y(t) = \cos(x(t)) + \cos(y(t-1))$$

when the input is Gaussian noise, and the coherence and phase spectra prove that F partly identifies the system. The linear model, based on $x(t)$, $y(t-1)$, and odd powers to 5, does not identify the filter, either qualitatively on the first row graph, or as measured by coherence and phase.

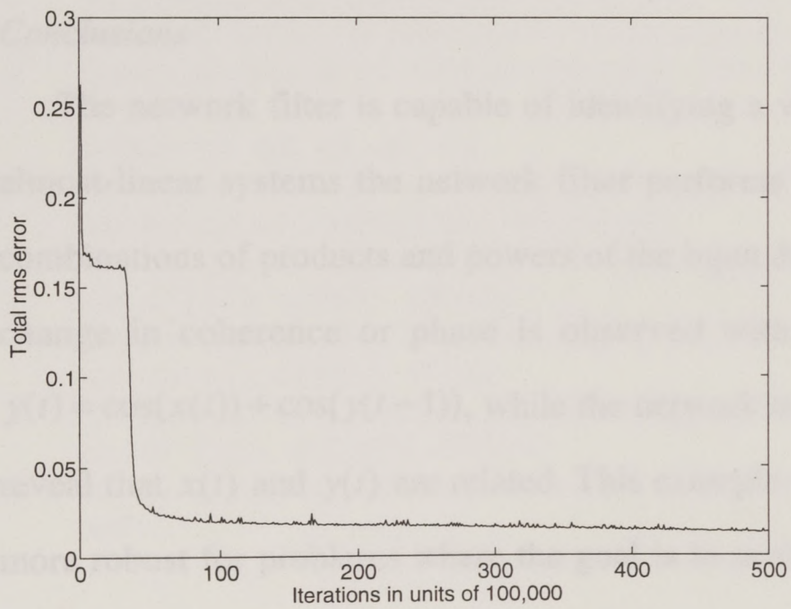


Figure 4.3 The total rms error between $x(t)$ and $y(t)$ during design of F. After several million iterations the error ceases to decrease, but shortly thereafter drops precipitously. The gradient method presented is not optimal, and provides no information as when to cease iteration.

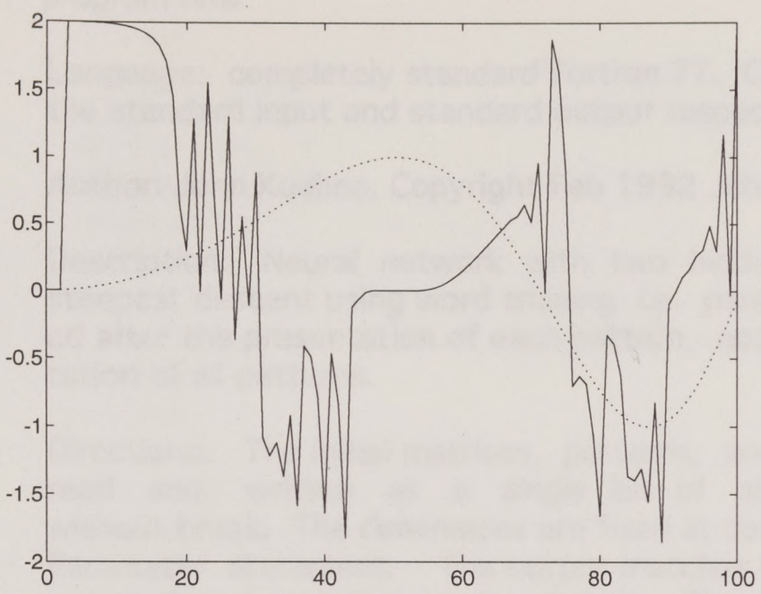


Figure 4.4 The first couple of cycles of the frequency sweep test illustrate why this filter is difficult to identify. At first filter output is smooth and correlated like the input, but it develops instabilities part-way through the first sinusoid cycle before smoothing out. These instabilities are set off by minute changes in the slope of $x(t)$. Phase space diagrams for a number of sinusoids show the instabilities are most complicated for low frequency sinusoids, and collapse into simpler states for higher frequencies.

Conclusions

The network filter is capable of identifying a variety of non-linear systems. For almost-linear systems the network filter performs at least as well as finding linear combinations of products and powers of the input data $z_j(t)$. However, no significant change in coherence or phase is observed with the linear model for the filter $y(t) = \cos(x(t)) + \cos(y(t-1))$, while the network model coherence and phase spectra reveal that $x(t)$ and $y(t)$ are related. This example suggests that the network filter is more robust for problems where the goal is to ascertain whether two time series are functionally related.

The main disadvantage of the network filter is the slow algorithm for finding the network parameters. Faster algorithms are available, even for the steepest descent procedure. A sensible strategy in system identification is to try the linear method first before pursuing the network solution, especially to give some insight about the form of $z_i(t)$.

```
      program nnw

c     Language: completely standard Fortran 77.  Channels 5 and 6  are
c     the standard input and standard output respectively.

c     Author: John Kuehne. Copyright Feb 1992 John Kuehne

c     Description:  Neural network with two hidden layers, quasi
c     steepest descent using word training  i.e.  parameters are updat-
c     ed after the presentation of each pattern,  not after the presen-
c     tation of all patterns.

c     Directions:  The initial matrices,  patterns,  and  targets  are
c     read  and  written  as  a  single  list  of  ascii numbers
c     without break.  The dimensions are fixed at compile time via  the
c     Parameter  statement.   The output matches the format of the in-
c     put so the program is easy  to  restart.  The  idea  is  to  use
c     Matlab  as  a  front-end  to  setup  the problem and examine the
c     results. Re: q1 input lines, q2  first  layer  nodes,  q3  second
c     layer nodes, q4 output lines, q5 patterns.  For reading and writ-
c     ing the list of numbers,   matrices are stored as the consecutive
c     column vectors,  followed by  the patterns which consist of  con-
c     secutive lists of  q5 numbers for  each of  the q1  input  lines,
c     followed  lastly  by the targets  which are written like the pat-
c     terns  i.e. lists of q5 numbers for each of the q4 output lines.

      integer q1, q2, q3, q4, q5
      parameter (q1 = 10, q2 = 40, q3 = 20, q4 = 2, q5 = 99)
      real w1(q2,q1), w2(q3,q2), w3(q4,q3)
      real t1(q2), t2(q3), t3(q4)
      real o1(q2), o2(q3), o3(q4)
      real pattern(q1,q5), target(q4,q5)
      real d3(q4), d2(q3), d1(q2)
      real dw1(q2,q1), dw2(q3,q2), dw3(q4,q3)
      real dt1(q2), dt2(q3), dt3(q4)
      integer i, j, k, iter, pass1, pass2, s
      real n


c     Read initial matrices, pattern, and target from standard input.
      do 1 j = 1,q1
        do 2 i=1,q2
          read(5,*) w1(i,j)
2       continue
1     continue

      do 3 j = 1,q2
        do 4 i=1,q3
          read(5,*) w2(i,j)
```

```
4       continue
3     continue

      do 5 j = 1,q3
        do 6 i=1,q4
          read(5,*) w3(i,j)
6       continue
5     continue

      do 7 i=1,q2
        read(5,*) t1(i)
7     continue

      do 8 i=1,q3
        read(5,*) t2(i)
8     continue

      do 9 i=1,q4
        read(5,*) t3(i)
9     continue

      do 10 j = 1,q1
        do 11 i=1,q5
          read(5,*) pattern(j,i)
11      continue
10    continue

      do 12 j =1,q4
        do 13 i=1,q5
          read(5,*) target(j,i)
13      continue
12    continue

      iter = 50
      n = 0.1

      do 14 pass1 = 1,iter
      do 15 pass2 = 1,100000

c     select pattern
      s = nint(rand(0)*float(q5-1))+1

c     First matrix-vector multiply (w1*pattern) and sigmoidal transfer.
      do 20 j = 1,q2
        o1(j) = w1(j,1) * pattern(1,s)
        do 30 k = 2,q1
          o1(j) = o1(j) + (w1(j,k) * pattern(k,s))
30      continue
        o1(j) = 1.0 / (1.0 + exp(-1.0 * (o1(j) + t1(j))))
```

```
20    continue

c     Second matrix-vector multiply.  Same as above.
      do 50 j=1,q3
        o2(j) = w2(j,1) * o1(1)
        do 60 k = 2,q2
          o2(j) = o2(j) + (w2(j,k) * o1(k))
60      continue
        o2(j) = 1.0 / (1.0 + exp(-1.0 * (o2(j) + t2(j))))
50    continue

c     Third matrix-vector multiply.  Same as above.
      do 80 j = 1,q4
        o3(j) = w3(j,1) * o2(1)
        do 90 k = 2,q3
          o3(j) = o3(j) + (w3(j,k) * o2(k))
90      continue
        o3(j) = 1.0 / (1.0 + exp(-1.0 * (o3(j) + t3(j))))
80    continue

c     Backpropogate errors.
      do 110 j = 1,q4
        d3(j) = (target(j,s) - o3(j)) * (o3(j)) *
     $          (1.0 - o3(j))
110   continue

      do 130 j = 1,q3
        d2(j) = w3(1,j) * d3(1)
        do 140 k = 2,q4
          d2(j) = d2(j) + (w3(k,j) * d3(k))
140     continue
      d2(j) = d2(j) * o2(j) * (1.0 - o2(j))
130   continue

      do 155 j = 1,q2
        d1(j) = w2(1,j) * d2(1)
        do 165 k = 2,q3
          d1(j) = d1(j) + (w2(k,j) * d2(k))
165     continue
        d1(j) =  d1(j) * o1(j) * (1.0 - o1(j))
155   continue

c     Estimate offset perturbations.
      do 170 j = 1,q2
        dt1(j) = 0.0
170   continue
        do 190 j = 1,q2
          dt1(j) = dt1(j) + d1(j)
190     continue
```

```
      do 200 j = 1,q2
        t1(j) = t1(j) + (dt1(j) * n)
200   continue

      do 210 j = 1,q3
        dt2(j) = 0.0
210   continue
        do 230 j = 1,q3
          dt2(j) = dt2(j) + d2(j)
230       continue
      do 240 j=1,q3
        t2(j) = t2(j) + (dt2(j) * n)
240   continue

      do 250 j = 1,q4
        dt3(j) = 0.0
250   continue
        do 270 j = 1,q4
          dt3(j) = dt3(j) + d3(j)
270       continue
      do 280 j = 1,q4
        t3(j) = t3(j) + (dt3(j) * n)
280   continue

c     Estimate weight perturbations for selected pattern.
      do 290 j = 1,q3
        do 300 i = 1,q4
          dw3(i,j) = d3(i) * o2(j)
300       continue
290   continue
      do 292 j = 1,q3
        do 293 i = 1,q4
          w3(i,j) = w3(i,j) + (dw3(i,j) * n)
293       continue
292   continue

      do 310 j = 1,q2
        do 320 i = 1,q3
          dw2(i,j) = d2(i) * o1(j)
320       continue
310   continue
      do 312 j = 1,q2
        do 313 i = 1,q3
          w2(i,j) = w2(i,j) + (dw2(i,j) * n)
313       continue
312   continue

      do 330 j = 1,q1
        do 340 i = 1,q2
```

```
            dw1(i,j) =  d1(i) * pattern(j,s)
340         continue
330       continue
        do 332 j = 1,q1
          do 333 i = 1,q2
            w1(i,j) = w1(i,j) + (dw1(i,j) * n)
333         continue
332       continue

15      continue
14      continue

c
c       Write final matrices to standard output as a single list.
c
        do 400 j = 1,q1
          do 405 i=1,q2
            write(6,*) w1(i,j)
405         continue
400       continue

        do 410 j = 1,q2
          do 415 i=1,q3
            write(6,*) w2(i,j)
415         continue
410       continue

        do 420 j = 1,q3
          do 425 i=1,q4
            write(6,*) w3(i,j)
425         continue
420       continue

        do 430 i=1,q2
          write(6,*) t1(i)
430       continue

        do 440 i=1,q3
          write(6,*) t2(i)
440       continue

        do 450 i=1,q4
          write(6,*) t3(i)
450       continue

c       write initial pattern and target as a single list
        do 460 j = 1,q1
          do 470 i=1,q5
            write(6,*) pattern(j,i)
```

```
470       continue
460     continue

      do 480 j =1,q4
        do 490 i=1,q5
          write(6,*) target(j,i)
490       continue
480     continue


      end
```

Appendix D

Annual component of polar motion

Throughout this dissertation emphasis has been placed on the continuous part of polar motion spectrum and special care has been taken to remove annual and semi-annual sinusoids. This section briefly describes the annual variation.

The annual motion of the pole is a prominent part of polar motion. It was correctly understood by early workers as arising partly by accumulation of cold dense air over the northern hemisphere each winter [*Munk and MacDonald*, 1960]. The importance of the annual and Chandler frequencies in the literature is partly an artifact of the old optical data, which are reliable only at frequencies near the resonant Chandler frequency of about +0.84 cpy (435d). Modern data are reliable at periods between days and decades and the emphasis of modern investigations is no longer centered on the annual line spectrum, which is described by just four numbers.

The annual component of polar motion is defined here as the 1 cpy line spectrum of polar motion. It can be described by the sinusoidal variation along the real and imaginary axes of the geographic reference frame

$$m(t) = \alpha \sin(\omega t + \phi_r) + i\beta \sin(\omega t + \phi_i) \qquad 1$$

which describes an ellipse, where t is time, ω is frequency, and the remaining constants specify magnitude and phase. Using the identity

$$\sin(x + y) = \cos(x)\sin(y) + \sin(x)\cos(y)$$

(1) expands to the linear combination

$$m(t) = a\cos(\omega t) + b\sin(\omega t) + i[c\cos(\omega t) + d\sin(\omega t)]. \qquad 2$$

The four parameters a,b,c,d can then be determined by the least squares projection of the polar motion time series onto $\cos(\omega t)$ and $\sin(\omega t)$ for the real and imaginary

components. However, this representation is tied to the rather arbitrary coordinate system of Greenwich. More geophysical meaning can be had by projecting the polar motion time series onto prograde (west-to-east) and retrograde (east-to-west) circular components, i.e.

$$Ae^{i\omega t} + Be^{-i\omega t}. \qquad 3$$

The complex constants $A = A_1 + iA_2$ and $B = B_1 + iB_2$ are called the prograde and retrograde vectors. The magnitudes of these vectors are the same regardless of the choice of 0 longitude, with $|A|+|B|$ the length of the semimajor axis of the ellipse, and $\|A|-|B\|$ the length of the semiminor axis. The relationship between the Cartesian parameters *a,b,c,d* and the complex quantities *A* and *B* is

$$A_1 = (a+d)/2, \quad A_2 = (c-b)/2 \quad B_1 = (a-d)/2 \quad B_2 = (b+c)/2. \qquad 4$$

The prograde and retrograde components are simply the coefficients of the discrete Fourier transform at ±1 cpy. The discussion above thus carries over to calculating the continuous part of the spectrum, producing the positive (prograde) and negative (retrograde) frequencies of spectral plots. The power spectrum, computed only from the magnitudes of the prograde and retrograde coefficients, is thus a coordinate-free representation.

Using the old optical data and global barometer data, a number of studies have shown a significant discrepancy at 1 cpy between polar motion and atmospheric variations with an inverted barometer ocean. The prograde atmospheric vectors of *Kuehne and Wilson* [1990] and *Wahr* [1982], both based on historical data, are about 100°, which is considerably west of polar motion (65°), while the magnitudes are slightly larger than polar motion. The emphasis on prograde vectors in these studies using optical data is a result of the amplification near the prograde resonant Chandler

wobble. The retrograde component, being much farther from the Chandler frequency, is more susceptible to noise.

The conclusion of *Kuehne and Wilson* [1990] is that there is another excitation source, and the results of the present study using the ECMWF data and space93 for the period 1985-1993 support that conclusion. The vectors for χ_t^{PL}, which is essentially an inverted barometer model, are comparable to those of *Wahr* [1982] and *Kuehne and Wilson* [1990]. Remarkably, the linear combination of land and SH ocean, described in Chapter 2, nearly matches the prograde and retrograde polar motion vectors in both phase and magnitude, further suggesting that the other excitation source is over the oceans, either as a non-inverted barometer response to pressure forcing, or some other effect of the winds.

Table 1

		Prograde		Retrograde	
Study		Phase	Magnitude	Phase	Magnitude
Wahr [1982] (1900-1958)		-100	10.2E-8	-94	10.0E-8
Kuehne and Wilson [1990] (1900-1985)		-100	8.6E-8	-94	7.9E-8
χ_t^{PL}	(1985-1993)	-101	6.0E-8	-106	6.7E-8
$\chi_t^{PO'}$	(1985-1993)	70	3.3E-8	-149	3.9E-8
χ_t^{A}	(1985-1993)	-80	5.3E-8	-110	6.5E-8
χ_t^{m} space93	(1985-1993)	-70	6.9E-8	-119	3.3E8
ILS	(1900-1985)	-64	7.7E-8	-141	9.4E-8

Note: phases are relative to January 1.

Bibliography

Barnes, R.T.H., R. Hide, A.A. White, and C.A. Wilson, Atmospheric angular momentum fluctuations, length-of-day changes and polar motion, *Proc. R. Soc. London Ser. A, 387*, 31-73, 1983.

Chao, B.F., A.Y. Au, Atmospheric excitation of the Earth's annual wobble: 1980-1988, *J. Geophys. Res., 96* (B4), 6577-6582, 1991.

Eubanks, T.M., J.A. Steppe, J.O. Dickey, R.D. Rosen, D.A. Salstein, Causes of rapid motions of the Earth's pole, *Nature*, *334* (6178), 115-119, 1988.

Eubanks, T.M., Variations in the orientation of the Earth, *Contributions of Space Geodesy to Geodynamics: Earth Dynamics, Geodyn. Ser.*, vol. 24, edited by D.E. Smith and D.L. Turcotte, pp. 1-54, AGU, Washington, D.C., 1993.

Gross, R.S., Correspondence between theory and observations of polar motion, *Geophys. J. Int., 109*, 162-170, 1992.

Gross, R.S., U.J. Lindqwister, Atmospheric excitation of polar motion during the GIG '91 measurement campaign, *Geophys. Res. Lett., 19* (9), 849-852, 1992.

International Earth Rotation Service, Annual report for 1993, Observatoire de Paris, Paris 1993.

Jeffreys, H., Causes contributory to the annual variation of latitude, *Mon. Not. R. Astron. Soc, 76*, 499-524, 1916.

Jeffreys, H., The variation of latitude, *Mon. Not. R. Astron. Soc., 100*, 139-154, 1940.

Jeffreys, H., The variation of latitude, *Mon. Not. R. Astron. Soc., 141,* 255-268, 1968.

Kuehne, J., C.R. Wilson, Terrestrial water storage and polar motion, *J. Geophys. Res.,* 96, 4337-4345, 1991.

Kuehne, J., C.R. Wilson, and S. Johnson, Atmospheric excitation of nonseasonal polar motion, *J. Geophys. Res., 98*, 19,973-19,978, 1993.

Lapedes, A., R. Farber, How neural networks work, Proceedings of IEEE, *Denver conference on neural nets 1987* (submitted). Los Alamos Nat. Lab. preprint LA-UR-88-418.

Lapedes, A., R. Farber, Nonlinear signal processing using neural networks: prediction and system modelling, Los Alamos National Laboratory, Lost Alamos, NM 87545.

Narendra, K.S., K. Parthasarathy, Identification and control of dynamical systems using neural networks, *IEEE Trans. on Neural Networks*, vol. 1, no. 1, March 1990.

Preisig, J.R., Polar motion, atmospheric angular momentum excitation and earthquakes - correlations and significance, *Geophys. J. Int, 108*, 161-178, 1992.

Tenorio, M., W. Lee, Self-organizing netowork for optimum supervised learning, *IEEE Trans. on Neural Networks,* vol. 1, no.1, March 1990.

Thompson, D.J., Spectrum estimation and harmonic analysis, *Proc. of the IEEE, 70* (9), 1055-1096, 1982.

Trenberth, K.E., J.G. Olson, An evaluation and intercomparison of global analyses from the National Meteorological Center and European Centre for Medium Range Weather Forecasts, *Bull. Am. Meteorol. Soc., 69* (9), 1047-1057, 1988.

Wilson, C.R., Discrete polar motion equations, *Geophys. J. R. Astron. Soc., 80*, 551-554, 1985.

Wilson, C.R., R.O. Vicente, An analysis of the homogeneous ILS polar motion series, *Geophys. J. R. Astron. Soc., 62*, 605-6161, 1980.

Wilson, C.R. and R.O. Vicente, Maximum likelihood estimates of polar motion parameters, *Variations in Earth Rotation*, *Geophys. Mono. Ser.*, vol. 59, edited by D.D. McCarthy and W.E. Carter, pp. 151-155, AGU, Washington, D.C., 1990.

The vita has been removed from the digitized version of this document.